

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Roger J. Davis et al.                      Art Unit : 1632  
Serial No. : 09/165,522                                      Examiner : M. Rao  
Filed : October 2, 1998  
Title : JNK3 MODULATORS AND METHODS OF USE

Commissioner for Patents  
Washington, D.C. 20231

TRANSMITTAL OF DUPLICATE COPY OF SEQUENCE LISTING

As requested by the Examiner on July 7, 2000, in a telephone conversation with J. Peter Fasse, a duplicate copy of the computer readable sequence listing, submitted on June 5, 2000, is enclosed.

Respectfully submitted,

Date: \_\_\_\_\_

*July 7, 2000*

*J. Peter Fasse*  
\_\_\_\_\_  
J. Peter Fasse  
Reg. No. 32,983

Fish & Richardson P.C.  
225 Franklin Street  
Boston, MA 02110-2804  
Telephone: (617) 542-5070  
Facsimile: (617) 542-8906

20068526.doc

RECEIVED  
JUL 12 2000  
TECH CENTER 1600/2900

CERTIFICATE OF MAILING BY FEDERAL EXPRESS

I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as Federal Express to Addressee with sufficient postage on the date indicated below and is addressed to Examiner Rao, USPTO, 1911 S. Clark Street, Arlington, VA 22202, Crystal Mall 1, Floor 7 Receptionist

\_\_\_\_\_  
Date of Deposit                      July 7, 2000

\_\_\_\_\_  
Signature

*Darlene J. Morin*

\_\_\_\_\_  
Darlene J. Morin  
Typed or Printed Name of Person Signing Certificate

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Davis, Roger J.  
Flavell, Richard A.  
Rakic, Pasko  
Whitmarsh, Alan  
Kuan, Chia-Yi  
Yang, Di

(ii) TITLE OF INVENTION: JNK3 MODULATORS AND METHODS OF USE

(iii) NUMBER OF SEQUENCES: 24

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Fish & Richardson P.C.  
(B) STREET: 225 Franklin Street  
(C) CITY: Boston  
(D) STATE: MA  
(E) COUNTRY: USA  
(F) ZIP: 02110-2804

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: Windows 95  
(D) SOFTWARE: FastSEQ for Windows Version 2.0b

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/165,522  
(B) FILING DATE: 02-OCT-1998

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 60/060,995  
(B) FILING DATE: 03-OCT-1997

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Fasse, J. Peter  
(B) REGISTRATION NUMBER: 32,983  
(C) REFERENCE/DOCKET NUMBER: 10363/005001

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 617/542-5070  
(B) TELEFAX: 617/542-8906  
(C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1505 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence  
(B) LOCATION: 68...1459

Part of  
12/13

GOV. 10 OCT 9:43

RECEIVED  
NOV 07/2000

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTATGCAAGA AACTGTTGAA TTAGACCCGT TTCCTATAGA TGAGAAACCA TACAAGCTGT	60
GGTATTT ATG AGC CTC CAT TTC TTA TAC TAC TGC AGT GAA CCA ACA TTG	109
Met Ser Leu His Phe Leu Tyr Tyr Cys Ser Glu Pro Thr Leu	
1 5 10	
GAT GTG AAA ATT GCC TTT TGT CAG GGA TTC GAT AAA CAA GTG GAT GTG	157
Asp Val Lys Ile Ala Phe Cys Gln Gly Phe Asp Lys Gln Val Asp Val	
15 20 25 30	
TCA TAT ATT GCC AAA CAT TAC AAC ATG AGC AAA AGC AAA GTT GAC AAC	205
Ser Tyr Ile Ala Lys His Tyr Asn Met Ser Lys Ser Lys Val Asp Asn	RECEIVED
35 40 45	
CAG TTC TAC AGT GTG GAA GTG GGA GAC TCA ACC TTC ACA GTT CTC AAG	253 12 200
Gln Phe Tyr Ser Val Glu Val Gly Asp Ser Thr Phe Thr Val Leu Lys	123 12 200
50 55 60	
CGC TAC CAG AAT CTA AAG CCT ATT GGC TCT GGG GCT CAG GGC ATA GTT	301
Arg Tyr Gln Asn Leu Lys Pro Ile Gly Ser Gly Ala Gln Gly Ile Val	
65 70 75	
TGT GCC GCG TAT GAT GCT GTC CTT GAC AGA AAT GTG GCC ATT AAG AAG	349
Cys Ala Ala Tyr Asp Ala Val Leu Asp Arg Asn Val Ala Ile Lys Lys	
80 85 90	
CTC AGC AGA CCC TTT CAG AAC CAA ACA CAT GCC AAG AGA GCG TAC CGG	397
Leu Ser Arg Pro Phe Gln Asn Gln Thr His Ala Lys Arg Ala Tyr Arg	
95 100 105 110	
GAG CTG GTC CTC ATG AAG TGT GTG AAC CAT AAA AAC ATT ATT AGT TTA	445
Glu Leu Val Leu Met Lys Cys Val Asn His Lys Asn Ile Ile Ser Leu	
115 120 125	
TTA AAT GTC TTC ACA CCC CAG AAA ACG CTG GAG GAG TTC CAA GAT GTT	493
Leu Asn Val Phe Thr Pro Gln Lys Thr Leu Glu Glu Phe Gln Asp Val	
130 135 140	
TAC TTA GTA ATG GAA CTG ATG GAT GCC AAC TTA TGT CAA GTG ATT CAG	541
Tyr Leu Val Met Glu Leu Met Asp Ala Asn Leu Cys Gln Val Ile Gln	
145 150 155	
ATG GAA TTA GAC CAT GAG CGA ATG TCT TAC CTG CTG TAC CAA ATG TTG	589
Met Glu Leu Asp His Glu Arg Met Ser Tyr Leu Tyr Gln Met Leu	
160 165 170	
TGT GGC ATT AAG CAC CTC CAT TCT GCT GGA ATT ATT CAC AGG GAT TTA	637
Cys Gly Ile Lys His Leu His Ser Ala Gly Ile Ile His Arg Asp Leu	
175 180 185 190	
AAA CCA AGT AAC ATT GTA GTC AAG TCT GAT TGC ACA TTG AAA ATC CTG	685
Lys Pro Ser Asn Ile Val Val Lys Ser Asp Cys Thr Leu Lys Ile Leu	
195 200 205	
GAC TTT GGA CTG GCC AGG ACA GCA GGC ACA AGC TTC ATG ATG ACT CCA	733
Asp Phe Gly Leu Ala Arg Thr Ala Gly Thr Ser Phe Met Met Thr Pro	
210 215 220	
TAT GTG GTG ACA CGT TAT TAC AGA GCC CCT GAG GTC ATC CTG GGG ATG	781
Tyr Val Val Thr Arg Tyr Tyr Arg Ala Pro Glu Val Ile Leu Gly Met	
225 230 235	

GGC TAC AAG GAG AAC GTG GAT ATA TGG TCT GTG GGA TGC ATT ATG GGA Gly Tyr Lys Glu Asn Val Asp Ile Trp Ser Val Gly Cys Ile Met Gly 240 245 250	829
GAA ATG GTT CGC CAC AAA ATC CTC TTT CCA GGA AGG GAC TAT ATT GAC Glu Met Val Arg His Lys Ile Leu Phe Pro Gly Arg Asp Tyr Ile Asp 255 260 265 270	877
CAG TGG AAT AAG GTA ATT GAA CAA CTA GGA ACA CCA TGT CCA GAA TTC Gln Trp Asn Lys Val Ile Glu Gln Leu Gly Thr Pro Cys Pro Glu Phe 275 280 285	925
ATG AAG AAA TTG CAA CCC ACA GTA AGA AAC TAT GTG GAG AAT CGG CCC Met Lys Lys Leu Gln Pro Thr Val Arg Asn Tyr Val Glu Asn Arg Pro 290 295 300	973
AAG TAT GCG GGA CTC ACC TTC CCC AAA CTC TTC CCA GAT TCC CTC TTC Lys Tyr Ala Gly Leu Thr Phe Pro Lys Leu Phe Pro Asp Ser Leu Phe 305 310 315	1021
CCA GCG GAC TCC GAG CAC AAT AAA CTC AAA GCC AGC CAA GCC AGG GAC Pro Ala Asp Ser Glu His Asn Lys Leu Lys Ala Ser Gln Ala Arg Asp 320 325 330	1069
TTG TTG TCA AAG ATG CTA GTG ATT GAC CCA GCA AAA AGA ATA TCA GTG Leu Leu Ser Lys Met Leu Val Ile Asp Pro Ala Lys Arg Ile Ser Val 335 340 345 350	1117
GAC GAC GCC TTA CAG CAT CCC TAC ATC AAC GTC TGG TAT GAC CCA GCC Asp Asp Ala Leu Gln His Pro Tyr Ile Asn Val Trp Tyr Asp Pro Ala 355 360 365	1165
GAA GTG GAG GCG CCT CCA CCT CAG ATA TAT GAC AAG CAG TTG GAT GAA Glu Val Glu Ala Pro Pro Pro Gln Ile Tyr Asp Lys Gln Leu Asp Glu 370 375 380	1213
AGA GAA CAC ACA ATT GAA GAA TGG AAA GAA CTT ATC TAC AAG GAA GTA Arg Glu His Thr Ile Glu Glu Trp Lys Glu Leu Ile Tyr Lys Glu Val 385 390 395	1261
ATG AAT TCA GAA GAA AAG ACT AAA AAT GGT GTA GTA AAA GGA CAG CCT Met Asn Ser Glu Glu Lys Thr Lys Asn Gly Val Val Lys Gly Gln Pro 400 405 410	1309
TCT CCT TCA GGT GCA GCA GTG AAC AGC AGT GAG AGT CTC CCT CCA TCC Ser Pro Ser Gly Ala Ala Val Asn Ser Ser Glu Ser Leu Pro Pro Ser 415 420 425 430	1357
TCG TCT GTC AAT GAC ATC TCC TCC ATG TCC ACC GAC CAG ACC CTG GCA Ser Ser Val Asn Asp Ile Ser Ser Met Ser Thr Asp Gln Thr Leu Ala 435 440 445	1405
TCT GAC ACT GAC AGC AGC CTG GAA GCC TCG GCA GGA CCC CTG GGT TGT Ser Asp Thr Asp Ser Ser Leu Glu Ala Ser Ala Gly Pro Leu Gly Cys 450 455 460	1453
TGC AGG TGACTAGCCG CCTGCCTGCG AAACCCAGCG TTCTTCAGGA GATGAT Cys Arg	1505

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 464 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Ser	Leu	His	Phe	Leu	Tyr	Tyr	Cys	Ser	Glu	Pro	Thr	Leu	Asp	Val
1				5					10					15	
Lys	Ile	Ala	Phe	Cys	Gln	Gly	Phe	Asp	Lys	Gln	Val	Asp	Val	Ser	Tyr
			20					25					30		
Ile	Ala	Lys	His	Tyr	Asn	Met	Ser	Lys	Ser	Lys	Val	Asp	Asn	Gln	Phe
		35					40					45			
Tyr	Ser	Val	Glu	Val	Gly	Asp	Ser	Thr	Phe	Thr	Val	Leu	Lys	Arg	Tyr
	50					55					60				
Gln	Asn	Leu	Lys	Pro	Ile	Gly	Ser	Gly	Ala	Gln	Gly	Ile	Val	Cys	Ala
65					70					75					80
Ala	Tyr	Asp	Ala	Val	Leu	Asp	Arg	Asn	Val	Ala	Ile	Lys	Lys	Leu	Ser
			85						90					95	
Arg	Pro	Phe	Gln	Asn	Gln	Thr	His	Ala	Lys	Arg	Ala	Tyr	Arg	Glu	Leu
			100					105					110		
Val	Leu	Met	Lys	Cys	Val	Asn	His	Lys	Asn	Ile	Ile	Ser	Leu	Leu	Asn
		115					120					125			
Val	Phe	Thr	Pro	Gln	Lys	Thr	Leu	Glu	Glu	Phe	Gln	Asp	Val	Tyr	Leu
	130					135					140				
Val	Met	Glu	Leu	Met	Asp	Ala	Asn	Leu	Cys	Gln	Val	Ile	Gln	Met	Glu
145					150					155					160
Leu	Asp	His	Glu	Arg	Met	Ser	Tyr	Leu	Leu	Tyr	Gln	Met	Leu	Cys	Gly
			165						170					175	
Ile	Lys	His	Leu	His	Ser	Ala	Gly	Ile	Ile	His	Arg	Asp	Leu	Lys	Pro
			180					185					190		
Ser	Asn	Ile	Val	Val	Lys	Ser	Asp	Cys	Thr	Leu	Lys	Ile	Leu	Asp	Phe
		195					200					205			
Gly	Leu	Ala	Arg	Thr	Ala	Gly	Thr	Ser	Phe	Met	Met	Thr	Pro	Tyr	Val
	210					215					220				
Val	Thr	Arg	Tyr	Tyr	Arg	Ala	Pro	Glu	Val	Ile	Leu	Gly	Met	Gly	Tyr
225					230					235					240
Lys	Glu	Asn	Val	Asp	Ile	Trp	Ser	Val	Gly	Cys	Ile	Met	Gly	Glu	Met
			245						250				255		
Val	Arg	His	Lys	Ile	Leu	Phe	Pro	Gly	Arg	Asp	Tyr	Ile	Asp	Gln	Trp
			260					265					270		
Asn	Lys	Val	Ile	Glu	Gln	Leu	Gly	Thr	Pro	Cys	Pro	Glu	Phe	Met	Lys
		275					280					285			
Lys	Leu	Gln	Pro	Thr	Val	Arg	Asn	Tyr	Val	Glu	Asn	Arg	Pro	Lys	Tyr
		290				295					300				
Ala	Gly	Leu	Thr	Phe	Pro	Lys	Leu	Phe	Pro	Asp	Ser	Leu	Phe	Pro	Ala
305					310					315					320
Asp	Ser	Glu	His	Asn	Lys	Leu	Lys	Ala	Ser	Gln	Ala	Arg	Asp	Leu	Leu
			325						330					335	
Ser	Lys	Met	Leu	Val	Ile	Asp	Pro	Ala	Lys	Arg	Ile	Ser	Val	Asp	Asp
		340						345					350		
Ala	Leu	Gln	His	Pro	Tyr	Ile	Asn	Val	Trp	Tyr	Asp	Pro	Ala	Glu	Val
		355					360					365			
Glu	Ala	Pro	Pro	Pro	Gln	Ile	Tyr	Asp	Lys	Gln	Leu	Asp	Glu	Arg	Glu
	370					375						380			

His Thr Ile Glu Glu Trp Lys Glu Leu Ile Tyr Lys Glu Val Met Asn  
 385 390 395 400  
 Ser Glu Glu Lys Thr Lys Asn Gly Val Val Lys Gly Gln Pro Ser Pro  
 405 410 415  
 Ser Gly Ala Ala Val Asn Ser Ser Glu Ser Leu Pro Pro Ser Ser  
 420 425 430  
 Val Asn Asp Ile Ser Ser Met Ser Thr Asp Gln Thr Leu Ala Ser Asp  
 435 440 445  
 Thr Asp Ser Ser Leu Glu Ala Ser Ala Gly Pro Leu Gly Cys Cys Arg  
 450 455 460

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2367 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAGAAATGGC	GTGGCAGGGG	ACCCAGCGAG	CCCAGAGGGA	TTTTGCCGCT	GCTTCCTCTA	60
CCCCTGTATT	TCACGCAGCT	CTCTAAATTG	ACTCAGCTCC	AGGCTAGTGT	GAGAAACACC	120
AACAGCAGGC	CCATCTCAGA	TCTTCACTAT	GGCAACTTAT	GCAAGAAACT	GTTGAATTAG	180
ACCCGTTTCC	TATAGATGAG	AAACCATACA	AGCTGTGGTA	TTTATGAGCC	TCCATTTCTT	240
ATACTACTGC	AGTGAACCAA	CATTGGATGT	GAAAATTGCC	TTTTGTCCAGG	GATTCGATAA	300
ACAAGTGGAT	GTGTCAATATA	TTGCCAAACA	TTACAACATG	AGCAAAAGCA	AAGTTGACAA	360
CCAGTTTCTAC	AGTGTGGAAG	TGGGAGACTC	AACCTTCACA	GTTCTCAAGC	GCTACCAGAA	420
TCTAAAGCCT	ATTGGCTCTG	GGGCTCAGGG	CATAGTTTGT	GCCGCGTATG	ATGCTGTCCT	480
TGACAGAAAT	GTGGCCATTA	AGAAGCTCAG	CAGACCCTTT	CAGAACCCAA	CACATGCCAA	540
GAGAGCGTAC	CGGGAGCTGG	TCCTCATGAA	GTGTGTGAAC	CATAAAAACA	TTATTAGTTT	600
ATTAATGTCT	TTACACCCCC	AGAAAACGCT	GGAGGAGTTC	CAAGATGTTT	ACTTAGTAAT	660
GGAACTGATG	GATGCCAACT	TATGTCAAGT	GATTCAGATG	GAATTAGACC	ATGAGCGAAT	720
GTCTTACCTG	CTGTACCAA	TGTTGTGTGG	CATTAAAGCAC	CTCCATTCTG	CTGGAATTAT	780
TCACAGGGAT	TTAAAACCAA	GTAACATTGT	AGTCAAGTCT	GATTGCACAT	TGAAAATCCT	840
GGACTTTGGA	CTGGCCAGGA	CAGCAGGCAC	AAGCTTCATG	ATGACTCCAT	ATGTGGTGAC	900
ACGTTATTAC	AGAGCCCCCTG	AGGTCATCCT	GGGGATGGGG	TACAAGGAGA	ACGTGGATAT	960
ATGGTCTGTG	GGATGCATTA	TGGGAGAAAT	GGTTCGCCAC	AAAATCCTCT	TTCCAGGAAG	1020
GGACTATATT	GACCAGTGGA	ATAAGGTAAT	TGAACAACTA	GGAACACCAT	GTCCAGAATT	1080
CATGAAGAAA	TTGCAACCCA	CAGTAAGAAA	CTATGTGGAG	AATCGGCCCA	AGTATGCGGG	1140
ACTCACCTTC	CCCAAACCTCT	TCCCAGATTG	CCTCTTCCCA	GCGGACTCCG	AGCACAAATA	1200
ACTCAAAGCC	AGCCAAGCCA	GGGACTTGTT	GTCAAAGATG	CTAGTGATTG	ACCCAGCAAA	1260
AAGAATATCA	GTGGACGACG	CCTTACAGCA	TCCCTACATC	AACGTCTGGT	ATGACCCAGC	1320
CGAAGTGGAG	GCGCCTCCAC	CTCAGATATA	TGACAAGCAG	TTGGATGAAA	GAGAACACAC	1380
AATTGAAGAA	TGGAAAGAAC	TTATCTACAA	GGAAGTAATG	AATTCAGAAG	AAAAGACTAA	1440
AAATGGTGTA	GTAAAAGGAC	AGCCTTCTCC	TTCAGGTGCA	GCAGTGAACA	GCAGTGAGAG	1500
TCTCCCTCCA	TCCTCGTCTG	TCAATGACAT	CTCCTCCATG	TCCACCGACC	AGACCCTGGC	1560
ATCTGACACT	GACAGCAGCC	TGGAAGCCTC	GGCAGGACCC	CTGGGTGTTT	GCAGGTGACT	1620
AGCCGCCTGC	CTGCGAAACC	CAGCGTTCTT	CAGGAGATGA	TGTGATGGAA	CACACACACA	1680
CGCAGACACA	CACACACACA	CAAATGCAGA	CACACAACAT	CAAGAAAACA	GCAAGGGAGA	1740
GAATCCAAGC	CTAAAATTAA	ATAAATCTTT	CAGCCTGCTT	CTTCCCCAGG	GTTCTGTATT	1800
GCAGCTAAGC	TCAAATGTAT	ATTAACTTC	TAGTTGCTCT	TGCTTTGGTC	TTCTTCCAAT	1860
GATGCTTACT	ACAGAAAGCA	AATCAGACAC	AATTAGAGAA	GCCTTTTCCA	TAAAGTGTA	1920
TTTTAATGGC	TGCAAAACCG	GCAACCTGTA	ACTGCCCTTT	TAAATGGCAT	GACAAGGTGT	1980
GCAGTGGCCC	CATCCAGCAT	GTGTGTGTCT	CTATCTTGCA	TCTACCTGCT	CCTTGGCCTA	2040
GTCAGATGGA	TGTAGATACA	GATCCGCATG	TGTCTGTATT	CATACAGCAC	TACTTACTTA	2100
GAGATGCTAC	TCTCAGTGTC	CTCAGGGCTC	TACCAAGACA	TAATGCACTG	GGGTACCACA	2160
TGGTCCATTT	CATGTGATCT	ATTACTCTGA	CATAAACCCA	TCTGTAATAT	ATTGCCAGTA	2220
TATAAGCTGT	TTAGTTTGTT	AATTGATTAA	ACTGTATGTC	TTATAAGAAA	ACATGTAAAG	2280

GGGGAATATA TTGGGGGAGT GAGCTCTCTC AGACCCTTGA AGATGTAGCT TCCAAATTTG  
AATGGATTAA ATGGCACCTG TATACCA

2340  
2367

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1773 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 92...1357

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATCTCAGATC TTCACTATGG CAACTTATGC AAGAACTGT TGAATTAGAC CCGTTTCCTA	60
TAGATGAGAA ACCATACAAG CTGTGGTATT T ATG AGC CTC CAT TTC TTA TAC	112
Met Ser Leu His Phe Leu Tyr	
1 5	
TAC TGC AGT GAA CCA ACA TTG GAT GTG AAA ATT GCC TTT TGT CAG GGA	160
Tyr Cys Ser Glu Pro Thr Leu Asp Val Lys Ile Ala Phe Cys Gln Gly	
10 15 20	
TTC GAT AAA CAA GTG GAT GTG TCA TAT ATT GCC AAA CAT TAC AAC ATG	208
Phe Asp Lys Gln Val Asp Val Ser Tyr Ile Ala Lys His Tyr Asn Met	
25 30 35	
AGC AAA AGC AAA GTT GAC AAC CAG TTC TAC AGT GTG GAA GTG GGA GAC	256
Ser Lys Ser Lys Val Asp Asn Gln Phe Tyr Ser Val Glu Val Gly Asp	
40 45 50 55	
TCA ACC TTC ACA GTT CTC AAG CGC TAC CAG AAT CTA AAG CCT ATT GGC	304
Ser Thr Phe Thr Val Leu Lys Arg Tyr Gln Asn Leu Lys Pro Ile Gly	
60 65 70	
TCT GGG GCT CAG GGC ATA GTT TGT GCC GCG TAT GAT GCT GTC CTT GAC	352
Ser Gly Ala Gln Gly Ile Val Cys Ala Ala Tyr Asp Ala Val Leu Asp	
75 80 85	
AGA AAT GTG GCC ATT AAG AAG CTC AGC AGA CCC TTT CAG AAC CAA ACA	400
Arg Asn Val Ala Ile Lys Lys Leu Ser Arg Pro Phe Gln Asn Gln Thr	
90 95 100	
CAT GCC AAG AGA GCG TAC CGG GAG CTG GTC CTC ATG AAG TGT GTG AAC	448
His Ala Lys Arg Ala Tyr Arg Glu Leu Val Leu Met Lys Cys Val Asn	
105 110 115	
CAT AAA AAC ATT ATT AGT TTA TTA AAT GTC TTC ACA CCC CAG AAA ACG	496
His Lys Asn Ile Ile Ser Leu Leu Asn Val Phe Thr Pro Gln Lys Thr	
120 125 130 135	
CTG GAG GAG TTC CAA GAT GTT TAC TTA GTA ATG GAA CTG ATG GAT GCC	544
Leu Glu Glu Phe Gln Asp Val Tyr Leu Val Met Glu Leu Met Asp Ala	
140 145 150	

AAC TTA TGT CAA GTG ATT CAG ATG GAA TTA GAC CAT GAG CGA ATG TCT Asn Leu Cys Gln Val Ile Gln Met Glu Leu Asp His Glu Arg Met Ser 155 160 165	592
TAC CTG CTG TAC CAA ATG TTG TGT GGC ATT AAG CAC CTC CAT TCT GCT Tyr Leu Leu Tyr Gln Met Leu Cys Gly Ile Lys His Leu His Ser Ala 170 175 180	640
GGA ATT ATT CAC AGG GAT TTA AAA CCA AGT AAC ATT GTA GTC AAG TCT Gly Ile Ile His Arg Asp Leu Lys Pro Ser Asn Ile Val Val Lys Ser 185 190 195	688
GAT TGC ACA TTG AAA ATC CTG GAC TTT GGA CTG GCC AGG ACA GCA GGC Asp Cys Thr Leu Lys Ile Leu Asp Phe Gly Leu Ala Arg Thr Ala Gly 200 205 210 215	736
ACA AGC TTC ATG ATG ACT CCA TAT GTG GTG ACA CGT TAT TAC AGA GCC Thr Ser Phe Met Met Thr Pro Tyr Val Val Thr Arg Tyr Tyr Arg Ala 220 225 230	784
CCT GAG GTC ATC CTG GGG ATG GGC TAC AAG GAG AAC GTG GAT ATA TGG Pro Glu Val Ile Leu Gly Met Gly Tyr Lys Glu Asn Val Asp Ile Trp 235 240 245	832
TCT GTG GGA TGC ATT ATG GGA GAA ATG GTT CGC CAC AAA ATC CTC TTT Ser Val Gly Cys Ile Met Gly Glu Met Val Arg His Lys Ile Leu Phe 250 255 260	880
CCA GGA AGG GAC TAT ATT GAC CAG TGG AAT AAG GTA ATT GAA CAA CTA Pro Gly Arg Asp Tyr Ile Asp Gln Trp Asn Lys Val Ile Glu Gln Leu 265 270 275	928
GGA ACA CCA TGT CCA GAA TTC ATG AAG AAA TTG CAA CCC ACA GTA AGA Gly Thr Pro Cys Pro Glu Phe Met Lys Lys Leu Gln Pro Thr Val Arg 280 285 290 295	976
AAC TAT GTG GAG AAT CGG CCC AAG TAT GCG GGA CTC ACC TTC CCC AAA Asn Tyr Val Glu Asn Arg Pro Lys Tyr Ala Gly Leu Thr Phe Pro Lys 300 305 310	1024
CTC TTC CCA GAT TCC CTC TTC CCA GCG GAC TCC GAG CAC AAT AAA CTC Leu Phe Pro Asp Ser Leu Phe Pro Ala Asp Ser Glu His Asn Lys Leu 315 320 325	1072
AAA GCC AGC CAA GCC AGG GAC TTG TTG TCA AAG ATG CTA GTG ATT GAC Lys Ala Ser Gln Ala Arg Asp Leu Leu Ser Lys Met Leu Val Ile Asp 330 335 340	1120
CCA GCA AAA AGA ATA TCA GTG GAC GAC GCC TTA CAG CAT CCC TAC ATC Pro Ala Lys Arg Ile Ser Val Asp Asp Ala Leu Gln His Pro Tyr Ile 345 350 355	1168
AAC GTC TGG TAT GAC CCA GCC GAA GTG GAG GCG CCT CCA CCT CAG ATA Asn Val Trp Tyr Asp Pro Ala Glu Val Glu Ala Pro Pro Pro Gln Ile 360 365 370 375	1216
TAT GAC AAG CAG TTG GAT GAA AGA GAA CAC ACA ATT GAA GAA TGG AAA Tyr Asp Lys Gln Leu Asp Glu Arg Glu His Thr Ile Glu Glu Trp Lys 380 385 390	1264



GAA CTT ATC TAC AAG GAA GTA ATG AAT TCA GAA GAA AAG ACT AAA AAT	1312
Glu Leu Ile Tyr Lys Glu Val Met Asn Ser Glu Glu Lys Thr Lys Asn	
395 400 405	
GGT GTA GTA AAA GGA CAG CCT TCT CCT TCA GCA CAG GTG CAG CAG	1357
Gly Val Val Lys Gly Gln Pro Ser Pro Ser Ala Gln Val Gln Gln	
410 415 420	
TGAACAGCAG TGAGAGTCTC CCTCCATCCT CGTCTGTCAA TGACATCTCC TCCATGTCCA	1417
CCGACCAGAC CCTGGCATCT GACACTGACA GCAGCCTGGA AGCCTCGGCA GGACCCCTGG	1477
GTTGTTGCAG GTGACTAGCC GCCTGCCTGC GAAACCCAGC GTTCTTCAGG AGATGATGTG	1537
ATGGAACACA CACACACGCA GACACACACA CACACACAAA TGCAGACACA CAACATCAAG	1597
AAAACAGCAA GGGAGAGAAT CCAAGCCTAA AATTAAATAA ATCTTTCAGC CTGCTTCTTC	1657
CCCAGGGTTC TGTATTGCAG CTAAGCTCAA ATGTATATTT AACTTCTAGT TGCTCTTGCT	1717
TTGGTCTTCT TCCAATGATG CTTACTACAG AAAGCAAATC AGACACAATT AGAGAA	1773

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ser	Leu	His	Phe	Leu	Tyr	Tyr	Cys	Ser	Glu	Pro	Thr	Leu	Asp	Val
1				5				10						15	
Lys	Ile	Ala	Phe	Cys	Gln	Gly	Phe	Asp	Lys	Gln	Val	Asp	Val	Ser	Tyr
		20					25						30		
Ile	Ala	Lys	His	Tyr	Asn	Met	Ser	Lys	Ser	Lys	Val	Asp	Asn	Gln	Phe
		35					40					45			
Tyr	Ser	Val	Glu	Val	Gly	Asp	Ser	Thr	Phe	Thr	Val	Leu	Lys	Arg	Tyr
	50					55				60					
Gln	Asn	Leu	Lys	Pro	Ile	Gly	Ser	Gly	Ala	Gln	Gly	Ile	Val	Cys	Ala
	65				70					75				80	
Ala	Tyr	Asp	Ala	Val	Leu	Asp	Arg	Asn	Val	Ala	Ile	Lys	Lys	Leu	Ser
			85					90						95	
Arg	Pro	Phe	Gln	Asn	Gln	Thr	His	Ala	Lys	Arg	Ala	Tyr	Arg	Glu	Leu
		100						105					110		
Val	Leu	Met	Lys	Cys	Val	Asn	His	Lys	Asn	Ile	Ile	Ser	Leu	Leu	Asn
	115					120						125			
Val	Phe	Thr	Pro	Gln	Lys	Thr	Leu	Glu	Glu	Phe	Gln	Asp	Val	Tyr	Leu
	130				135					140					
Val	Met	Glu	Leu	Met	Asp	Ala	Asn	Leu	Cys	Gln	Val	Ile	Gln	Met	Glu
	145				150				155					160	
Leu	Asp	His	Glu	Arg	Met	Ser	Tyr	Leu	Leu	Tyr	Gln	Met	Leu	Cys	Gly
			165					170					175		
Ile	Lys	His	Leu	His	Ser	Ala	Gly	Ile	Ile	His	Arg	Asp	Leu	Lys	Pro
		180					185						190		
Ser	Asn	Ile	Val	Val	Lys	Ser	Asp	Cys	Thr	Leu	Lys	Ile	Leu	Asp	Phe
	195					200						205			
Gly	Leu	Ala	Arg	Thr	Ala	Gly	Thr	Ser	Phe	Met	Met	Thr	Pro	Tyr	Val
	210				215					220					
Val	Thr	Arg	Tyr	Tyr	Arg	Ala	Pro	Glu	Val	Ile	Leu	Gly	Met	Gly	Tyr
	225				230					235				240	
Lys	Glu	Asn	Val	Asp	Ile	Trp	Ser	Val	Gly	Cys	Ile	Met	Gly	Glu	Met
			245					250					255		

Val Arg His Lys Ile Leu Phe Pro Gly Arg Asp Tyr Ile Asp Gln Trp  
260 265 270  
Asn Lys Val Ile Glu Gln Leu Gly Thr Pro Cys Pro Glu Phe Met Lys  
275 280 285  
Lys Leu Gln Pro Thr Val Arg Asn Tyr Val Glu Asn Arg Pro Lys Tyr  
290 295 300  
Ala Gly Leu Thr Phe Pro Lys Leu Phe Pro Asp Ser Leu Phe Pro Ala  
305 310 315 320  
Asp Ser Glu His Asn Lys Leu Lys Ala Ser Gln Ala Arg Asp Leu Leu  
325 330 335  
Ser Lys Met Leu Val Ile Asp Pro Ala Lys Arg Ile Ser Val Asp Asp  
340 345 350  
Ala Leu Gln His Pro Tyr Ile Asn Val Trp Tyr Asp Pro Ala Glu Val  
355 360 365  
Glu Ala Pro Pro Pro Gln Ile Tyr Asp Lys Gln Leu Asp Glu Arg Glu  
370 375 380  
His Thr Ile Glu Glu Trp Lys Glu Leu Ile Tyr Lys Glu Val Met Asn  
385 390 395 400  
Ser Glu Glu Lys Thr Lys Asn Gly Val Val Lys Gly Gln Pro Ser Pro  
405 410 415  
Ser Ala Gln Val Gln Gln  
420

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2372 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAGAAATGGC	GTGGCAGGGG	ACCCAGCGAG	CCCAGAGGGA	TTTTGCCGCT	GCTTCCTCTA	60
CCCCTGTATT	TCACGCAGCT	CTCTAAATTG	ACTCAGCTCC	AGGCTAGTGT	GAGAAACACC	120
AACAGCAGGC	CCATCTCAGA	TCTTCACTAT	GGCAACTTAT	GCAAGAACT	GTTGAATTAG	180
ACCCGTTTCC	TATAGATGAG	AAACCATACA	AGCTGTGGTA	TTTATGAGCC	TCCATTTCTT	240
ATACTACTGC	AGTGAACCAA	CATTGGATGT	GAAAATTGCC	TTTTGTCAGG	GATTCGATAA	300
ACAAGTGGAT	GTGTCATATA	TTGCCAAACA	TTACAACATG	AGCAAAAGCA	AAGTTGACAA	360
CCAGTTCTAC	AGTGTGGAAG	TGGGAGACTC	AACCTTCACA	GTTCTCAAGC	GCTACCAGAA	420
TCTAAAGCCT	ATTGGCTCTG	GGGCTCAGGG	CATAGTTTGT	GCCGCGTATG	ATGCTGTCCT	480
TGACAGAAAT	GTGGCCATTA	AGAAGCTCAG	CAGACCCTTT	CAGAACCAAA	CACATGCCAA	540
GAGAGCGTAC	CGGGAGCTGG	TCCTCATGAA	GTGTGTGAAC	CATAAAAACA	TTATTAGTTT	600
ATTAAATGTC	TTCACACCCC	AGAAAACGCT	GGAGGAGTTC	CAAGATGTTT	ACTTAGTAAT	660
GGAAGTGATG	GATGCCAACT	TATGTCAAGT	GATTCAGATG	GAATTAGACC	ATGAGCGAAT	720
GTCTTACCTG	CTGTACCAAA	TGTTGTGTGG	CATTAAGCAC	CTCCATTCTG	CTGGAATTAT	780
TCACAGGGAT	TTAAAACCAA	GTAACATTGT	AGTCAAGTCT	GATTGCACAT	TGAAATCCT	840
GGACTTTGGA	CTGGCCAGGA	CAGCAGGCAC	AAGCTTCATG	ATGACTCCAT	ATGTGGTGAC	900
ACGTTATTAC	AGAGCCCCTG	AGGTCATCCT	GGGGATGGGC	TACAAGGAGA	ACGTGGATAT	960
ATGGTCTGTG	GGATGCATTA	TGGGAGAAAT	GGTTCGCCAC	AAAATCCTCT	TTCCAGGAAG	1020
GGACTATATT	GACCAGTGGA	ATAAGGTAAT	TGAACAATA	GGAACACCAT	GTCCAGAATT	1080
CATGAAGAAA	TTGCAACCCA	CAGTAAGAAA	CTATGTGGAG	AATCGGCCCA	AGTATGCGGG	1140
ACTCACCTTC	CCCAAACTCT	TCCCAGATTG	CCTCTTCCCA	GCGGACTCCG	AGCACAATAA	1200
ACTCAAAGCC	AGCCAAGCCA	GGGACTTGTT	GTCAAAGATG	CTAGTGATTG	ACCCAGCAAA	1260
AAGAATATCA	GTGGACGACG	CCTTACAGCA	TCCCTACATC	AACGTCTGGT	ATGACCCAGC	1320
CGAAGTGGAG	GCGCCTCCAC	CTCAGATATA	TGACAAGCAG	TTGGATGAAA	GAGAACACAC	1380
AATTGAAGAA	TGGAAAGAAC	TTATCTACAA	GGAAGTAATG	AATTCAGAAG	AAAAGACTAA	1440
AAATGGTGTA	GTAAAAGGAC	AGCCTTCTCC	TTCAGCACAG	GTGCAGCAGT	GAACAGCAGT	1500
GAGAGTCTCC	CTCCATCCTC	GTCTGTCAAT	GACATCTCCT	CCATGTCCAC	CGACCAGACC	1560
CTGGCATCTG	ACACTGACAG	CAGCCTGGAA	GCCTCGGCAG	GACCCCTGGG	TTGTTGCAGG	1620

TGACTAGCCG	CCTGCCTGCG	AAACCCAGCG	TTCTTCAGGA	GATGATGTGA	TGGAACACAC	1680
ACACACGCAG	ACACACACAC	ACACACAAAT	GCAGACACAC	AACATCAAGA	AAACAGCAAG	1740
GGAGAGAATC	CAAGCCTAAA	ATTAAATAAA	TCTTTCAGCC	TGCTTCTTCC	CCAGGGTTCT	1800
GTATTGCAGC	TAAGCTCAAA	TGTATATTTA	ACTTCTAGTT	GCTCTTGCTT	TGGTCTTCTT	1860
CCAATGATGC	TTACTACAGA	AAGCAAATCA	GACACAATTA	GAGAAGCCTT	TTCCATAAAG	1920
TGTAATTTTA	ATGGCTGCAA	AACCGGCAAC	CTGTAACTGC	CCTTTTAAAT	GGCATGACAA	1980
GGTGTGCACT	GGCCCCATCC	AGCATGTGTG	TGTCTCTATC	TTGCATCTAC	CTGCTCCTTG	2040
GCCTAGTCAG	ATGGATGTAG	ATACAGATCC	GCATGTGTCT	GTATTCATAC	AGCACTACTT	2100
ACTTAGAGAT	GCTACTCTCA	GTGTCCTCAG	GGCTCTACCA	AGACATAATG	CACTGGGGTA	2160
CCACATGGTC	CATTTTCATGT	GATCTATTAC	TCTGACATAA	ACCCATCTGT	AATATATTGC	2220
CAGTATATAA	GCTGTTTAGT	TTGTTAATTG	ATTAAACTGT	ATGTCTTATA	AGAAAACATG	2280
TAAAGGGGGA	ATATATTGGG	GGAGTGAGCT	CTCTCAGACC	CTTGAAGATG	TAGCTTCCAA	2340
ATTTGAATGG	ATTAAATGGC	ACCTGTATAC	CA			2372

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2372 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 224...1489

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAGAAATGGC	GTGGCAGGGG	ACCCAGCGAG	CCCAGAGGGA	TTTGGCCGCT	GCTTCCTCTA	60
CCCCTGTATT	TCACGCAGCT	CTCTAAATTG	ACTCAGCTCC	AGGCTAGTGT	GAGAAACACC	120
AACAGCAGGC	CCATCTCAGA	TCTTCACTAT	GGCAACTTAT	GCAAGAAACT	GTTGAATTAG	180
ACCCGTTTCC	TATAGATGAG	AAACCATACA	AGCTGTGGTA	TTT ATG AGC CTC CAT		235
				Met Ser Leu His		
				1		
TTC TTA TAC TAC TGC AGT GAA CCA ACA TTG GAT GTG AAA ATT GCC TTT						283
Phe Leu Tyr Tyr Cys Ser Glu Pro Thr Leu Asp Val Lys Ile Ala Phe						
5 10 15 20						
TGT CAG GGA TTC GAT AAA CAA GTG GAT GTG TCA TAT ATT GCC AAA CAT						331
Cys Gln Gly Phe Asp Lys Gln Val Asp Val Ser Tyr Ile Ala Lys His						
25 30 35						
TAC AAC ATG AGC AAA AGC AAA GTT GAC AAC CAG TTC TAC AGT GTG GAA						379
Tyr Asn Met Ser Lys Ser Lys Val Asp Asn Gln Phe Tyr Ser Val Glu						
40 45 50						
GTG GGA GAC TCA ACC TTC ACA GTT CTC AAG CGC TAC CAG AAT CTA AAG						427
Val Gly Asp Ser Thr Phe Thr Val Leu Lys Arg Tyr Gln Asn Leu Lys						
55 60 65						
CCT ATT GGC TCT GGG GCT CAG GGC ATA GTT TGT GCC GCG TAT GAT GCT						475
Pro Ile Gly Ser Gly Ala Gln Gly Ile Val Cys Ala Ala Tyr Asp Ala						
70 75 80						
GTC CTT GAC AGA AAT GTG GCC ATT AAG AAG CTC AGC AGA CCC TTT CAG						523
Val Leu Asp Arg Asn Val Ala Ile Lys Lys Leu Ser Arg Pro Phe Gln						
85 90 95 100						

AAC CAA ACA CAT GCC AAG AGA GCG TAC CGG GAG CTG GTC CTC ATG AAG Asn Gln Thr His Ala Lys Arg Ala Tyr Arg Glu Leu Val Leu Met Lys 105 110 115	571
TGT GTG AAC CAT AAA AAC ATT ATT AGT TTA TTA AAT GTC TTC ACA CCC Cys Val Asn His Lys Asn Ile Ile Ser Leu Leu Asn Val Phe Thr Pro 120 125 130	619
CAG AAA ACG CTG GAG GAG TTC CAA GAT GTT TAC TTA GTA ATG GAA CTG Gln Lys Thr Leu Glu Glu Phe Gln Asp Val Tyr Leu Val Met Glu Leu 135 140 145	667
ATG GAT GCC AAC TTA TGT CAA GTG ATT CAG ATG GAA TTA GAC CAT GAG Met Asp Ala Asn Leu Cys Gln Val Ile Gln Met Glu Leu Asp His Glu 150 155 160	715
CGA ATG TCT TAC CTG CTG TAC CAA ATG TTG TGT GGC ATT AAG CAC CTC Arg Met Ser Tyr Leu Leu Tyr Gln Met Leu Cys Gly Ile Lys His Leu 165 170 175 180	763
CAT TCT GCT GGA ATT ATT CAC AGG GAT TTA AAA CCA AGT AAC ATT GTA His Ser Ala Gly Ile Ile His Arg Asp Leu Lys Pro Ser Asn Ile Val 185 190 195	811
GTC AAG TCT GAT TGC ACA TTG AAA ATC CTG GAC TTT GGA CTG GCC AGG Val Lys Ser Asp Cys Thr Leu Lys Ile Leu Asp Phe Gly Leu Ala Arg 200 205 210	859
ACA GCA GGC ACA AGC TTC ATG ATG ACT CCA TAT GTG GTG ACA CGT TAT Thr Ala Gly Thr Ser Phe Met Met Thr Pro Tyr Val Val Thr Arg Tyr 215 220 225	907
TAC AGA GCC CCT GAG GTC ATC CTG GGG ATG GGC TAC AAG GAG AAC GTG Tyr Arg Ala Pro Glu Val Ile Leu Gly Met Gly Tyr Lys Glu Asn Val 230 235 240	955
GAT ATA TGG TCT GTG GGA TGC ATT ATG GGA GAA ATG GTT CGC CAC AAA Asp Ile Trp Ser Val Gly Cys Ile Met Gly Glu Met Val Arg His Lys 245 250 255 260	1003
ATC CTC TTT CCA GGA AGG GAC TAT ATT GAC CAG TGG AAT AAG GTA ATT Ile Leu Phe Pro Gly Arg Asp Tyr Ile Asp Gln Trp Asn Lys Val Ile 265 270 275	1051
GAA CAA CTA GGA ACA CCA TGT CCA GAA TTC ATG AAG AAA TTG CAA CCC Glu Gln Leu Gly Thr Pro Cys Pro Glu Phe Met Lys Lys Leu Gln Pro 280 285 290	1099
ACA GTA AGA AAC TAT GTG GAG AAT CGG CCC AAG TAT GCG GGA CTC ACC Thr Val Arg Asn Tyr Val Glu Asn Arg Pro Lys Tyr Ala Gly Leu Thr 295 300 305	1147
TTC CCC AAA CTC TTC CCA GAT TCC CTC TTC CCA GCG GAC TCC GAG CAC Phe Pro Lys Leu Phe Pro Asp Ser Leu Phe Pro Ala Asp Ser Glu His 310 315 320	1195
AAT AAA CTC AAA GCC AGC CAA GCC AGG GAC TTG TTG TCA AAG ATG CTA Asn Lys Leu Lys Ala Ser Gln Ala Arg Asp Leu Leu Ser Lys Met Leu 325 330 335 340	1243

GTG ATT GAC CCA GCA AAA AGA ATA TCA GTG GAC GAC GCC TTA CAG CAT	1291
Val Ile Asp Pro Ala Lys Arg Ile Ser Val Asp Asp Ala Leu Gln His	
345 350 355	
CCC TAC ATC AAC GTC TGG TAT GAC CCA GCC GAA GTG GAG GCG CCT CCA	1339
Pro Tyr Ile Asn Val Trp Tyr Asp Pro Ala Glu Val Glu Ala Pro Pro	
360 365 370	
CCT CAG ATA TAT GAC AAG CAG TTG GAT GAA AGA GAA CAC ACA ATT GAA	1387
Pro Gln Ile Tyr Asp Lys Gln Leu Asp Glu Arg Glu His Thr Ile Glu	
375 380 385	
GAA TGG AAA GAA CTT ATC TAC AAG GAA GTA ATG AAT TCA GAA GAA AAG	1435
Glu Trp Lys Glu Leu Ile Tyr Lys Glu Val Met Asn Ser Glu Glu Lys	
390 395 400	
ACT AAA AAT GGT GTA GTA AAA GGA CAG CCT TCT CCT TCA GCA CAG GTG	1483
Thr Lys Asn Gly Val Val Lys Gly Gln Pro Ser Pro Ser Ala Gln Val	
405 410 415 420	
CAG CAG TGAACAGCAG TGAGAGTCTC CCTCCATCCT CGTCTGTCAA TGACATCTCC	1539
Gln Gln	

TCCATGTCCA CCGACCAGAC CCTGGCATCT GACACTGACA GCAGCCTGGA AGCCTCGGCA	1599
GGACCCCTGG GTTGTTCAG GTGACTAGCC GCCTGCCTGC GAAACCCAGC GTTCTTCAGG	1659
AGATGATGTG ATGGAACACA CACACACGCA GACACACACA CACACACAAA TGCAGACACA	1719
CAACATCAAG AAAACAGCAA GGGAGAGAAT CCAAGCCTAA AATTAAATAA ATCTTTCAGC	1779
CTGCTTCTTC CCCAGGGTTC TGTATTGCAG CTAAGCTCAA ATGTATATTT AACTTCTAGT	1839
TGCTCTTGCT TTGGTCTTCT TCCAATGATG CTTACTACAG AAAGCAAATC AGACACAATT	1899
AGAGAAGCCT TTTCCATAAA GTGTAATTTT AATGGCTGCA AAACCGGCAA CCTGTAACTG	1959
CCCTTTTAAA TGGCATGACA AGGTGTGCAG TGGCCCCATC CAGCATGTGT GTGTCTCTAT	2019
CTTGCATCTA CCTGCTCCTT GGCCTAGTCA GATGGATGTA GATACAGATC CGCATGTGTC	2079
TGTATTCATA CAGCACTACT TACTTAGAGA TGCTACTCTC AGTGTCTCTA GGGCTCTACC	2139
AAGACATAAT GCACTGGGGT ACCACATGGT CCATTTTCATG TGATCTATTA CTCTGACATA	2199
AACCCATCTG TAATATATTG CCAGTATATA AGCTGTTTAG TTTGTTAATT GATTAACTG	2259
TATGTCTTAT AAGAAAACAT GTAAAGGGGG AATATATTGG GGGAGTGAGC TCTCTCAGAC	2319
CCTTGAAGAT GTAGCTTCCA AATTTGAATG GATTAAATGG CACCTGTATA CCA	2372

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Ser Leu His Phe Leu Tyr Tyr Cys Ser Glu Pro Thr Leu Asp Val	
1 5 10 15	
Lys Ile Ala Phe Cys Gln Gly Phe Asp Lys Gln Val Asp Val Ser Tyr	
20 25 30	
Ile Ala Lys His Tyr Asn Met Ser Lys Ser Lys Val Asp Asn Gln Phe	
35 40 45	
Tyr Ser Val Glu Val Gly Asp Ser Thr Phe Thr Val Leu Lys Arg Tyr	
50 55 60	
Gln Asn Leu Lys Pro Ile Gly Ser Gly Ala Gln Gly Ile Val Cys Ala	
65 70 75 80	

Ala	Tyr	Asp	Ala	Val	Leu	Asp	Arg	Asn	Val	Ala	Ile	Lys	Lys	Leu	Ser	
				85					90					95		
Arg	Pro	Phe	Gln	Asn	Gln	Thr	His	Ala	Lys	Arg	Ala	Tyr	Arg	Glu	Leu	
			100					105					110			
Val	Leu	Met	Lys	Cys	Val	Asn	His	Lys	Asn	Ile	Ile	Ser	Leu	Leu	Asn	
		115				120						125				
Val	Phe	Thr	Pro	Gln	Lys	Thr	Leu	Glu	Glu	Phe	Gln	Asp	Val	Tyr	Leu	
	130					135					140					
Val	Met	Glu	Leu	Met	Asp	Ala	Asn	Leu	Cys	Gln	Val	Ile	Gln	Met	Glu	
	145				150				155						160	
Leu	Asp	His	Glu	Arg	Met	Ser	Tyr	Leu	Leu	Tyr	Gln	Met	Leu	Cys	Gly	
			165					170						175		
Ile	Lys	His	Leu	His	Ser	Ala	Gly	Ile	Ile	His	Arg	Asp	Leu	Lys	Pro	
			180					185					190			
Ser	Asn	Ile	Val	Val	Lys	Ser	Asp	Cys	Thr	Leu	Lys	Ile	Leu	Asp	Phe	
		195					200					205				
Gly	Leu	Ala	Arg	Thr	Ala	Gly	Thr	Ser	Phe	Met	Met	Thr	Pro	Tyr	Val	
	210					215					220					
Val	Thr	Arg	Tyr	Tyr	Arg	Ala	Pro	Glu	Val	Ile	Leu	Gly	Met	Gly	Tyr	
	225				230					235					240	
Lys	Glu	Asn	Val	Asp	Ile	Trp	Ser	Val	Gly	Cys	Ile	Met	Gly	Glu	Met	
			245					250						255		
Val	Arg	His	Lys	Ile	Leu	Phe	Pro	Gly	Arg	Asp	Tyr	Ile	Asp	Gln	Trp	
			260					265					270			
Asn	Lys	Val	Ile	Glu	Gln	Leu	Gly	Thr	Pro	Cys	Pro	Glu	Phe	Met	Lys	
		275					280					285				
Lys	Leu	Gln	Pro	Thr	Val	Arg	Asn	Tyr	Val	Glu	Asn	Arg	Pro	Lys	Tyr	
	290					295					300					
Ala	Gly	Leu	Thr	Phe	Pro	Lys	Leu	Phe	Pro	Asp	Ser	Leu	Phe	Pro	Ala	
	305				310					315					320	
Asp	Ser	Glu	His	Asn	Lys	Leu	Lys	Ala	Ser	Gln	Ala	Arg	Asp	Leu	Leu	
			325					330						335		
Ser	Lys	Met	Leu	Val	Ile	Asp	Pro	Ala	Lys	Arg	Ile	Ser	Val	Asp	Asp	
			340					345					350			
Ala	Leu	Gln	His	Pro	Tyr	Ile	Asn	Val	Trp	Tyr	Asp	Pro	Ala	Glu	Val	
		355					360					365				
Glu	Ala	Pro	Pro	Pro	Gln	Ile	Tyr	Asp	Lys	Gln	Leu	Asp	Glu	Arg	Glu	
	370					375					380					
His	Thr	Ile	Glu	Glu	Trp	Lys	Glu	Leu	Ile	Tyr	Lys	Glu	Val	Met	Asn	
	385				390					395					400	
Ser	Glu	Glu	Lys	Thr	Lys	Asn	Gly	Val	Val	Lys	Gly	Gln	Pro	Ser	Pro	
			405					410						415		
Ser	Ala	Gln	Val	Gln	Gln											
				420												

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1975 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 364...1641

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCCTCCTTAT TCCGGTTTGG AATGTGGCTA ATGAAAGCCC AGTAGGAGGA TTTCTGGGGC

60

AAACAGGTGG	ACCAGGATCC	TGGTTCTCAG	GCACGGAATG	GCTATTGTGA	GAGCGCCACC	120
AGCAGGACCA	TCGCAGATCT	TGGTTATGGC	TGCTCACGCA	AGAGGCTGTT	GATGTAGACC	180
CCCTTTCCCG	TAGATGAGAA	ATCACACGAG	CAGTGGTATT	TATGAGCCTC	CATTTCCTAT	240
ACTACTGCAG	TGAACCAACC	TTGGATGTGA	AAATTGCCTT	TTGTCAGGTG	TGTGTTCTTT	300
ACAGGTAAAA	CAAAGGGATT	CGACAAACAC	GTGGATGTGT	CTTCTGTTGT	CAAACATTAC	360
AAC ATG AGC AAA AGC AAG GTA GAT AAC CAG TTC	TAC AGT GTG GAA GTG	408				
Met Ser Lys Ser Lys Val Asp Asn Gln Phe	Tyr Ser Val Glu Val					
1	5	10	15			
GGA GAC TCA ACC TTC ACA GTT CTA AAG CGC TAC CAG AAC CTG AAG CCG	456					
Gly Asp Ser Thr Phe Thr Val Leu Lys Arg Tyr Gln Asn Leu Lys Pro						
20	25	30				
ATC GGC TCT GGG GCT CAG GGA ATA GTT TGT GCT GCG TAT GAC GCT GTC	504					
Ile Gly Ser Gly Ala Gln Gly Ile Val Cys Ala Ala Tyr Asp Ala Val						
35	40	45				
CTC GAC AGA AAT GTG GCC ATT AAG AAG CTC AGC AGA CCC TTC CAG AAC	552					
Leu Asp Arg Asn Val Ala Ile Lys Lys Leu Ser Arg Pro Phe Gln Asn						
50	55	60				
CAA ACT CAT GCC AAG AGG GCT TAC CGG GAG CTG GTC CTC ATG AAG TGT	600					
Gln Thr His Ala Lys Arg Ala Tyr Arg Glu Leu Val Leu Met Lys Cys						
65	70	75				
GTG AAC CAT AAA AAC ATT ATT AGC TTA TTA AAT GTC TTT ACA CCC CAG	648					
Val Asn His Lys Asn Ile Ile Ser Leu Leu Asn Val Phe Thr Pro Gln						
80	85	90				
AAA ACA CTG GAG GAG TTC CAA GAT GTT TAC TTA GTG ATG GAA CTG ATG	696					
Lys Thr Leu Glu Glu Phe Gln Asp Val Tyr Leu Val Met Glu Leu Met						
100	105	110				
GAC GCC AAC TTG TGT CAG GTG ATT CAG ATG GAG CTG GAC CAC GAG CGG	744					
Asp Ala Asn Leu Cys Gln Val Ile Gln Met Glu Leu Asp His Glu Arg						
115	120	125				
ATG TCG TAC TTG CTG TAC CAG ATG CTG TCG GCG ATC AAA CAC CTC CAC	792					
Met Ser Tyr Leu Leu Tyr Gln Met Leu Ser Ala Ile Lys His Leu His						
130	135	140				
TCC GCT GGG ATC ATC CAC AGG GAC TTA AAA CCC AGT AAC ATC GTA GTC	840					
Ser Ala Gly Ile Ile His Arg Asp Leu Lys Pro Ser Asn Ile Val Val						
145	150	155				
AAG TCT GAT TGC ACA CTG AAA ATC CTG GAC TTT GGA CTG GCC AGG ACA	888					
Lys Ser Asp Cys Thr Leu Lys Ile Leu Asp Phe Gly Leu Ala Arg Thr						
160	165	170				
GCG GGC ACA AGC TTC ATG ATG ACT CCG TAT GTG GTG ACG AGA TAT TAC	936					
Ala Gly Thr Ser Phe Met Met Thr Pro Tyr Val Val Thr Arg Tyr Tyr						
180	185	190				
AGA GCC CCC GAG GTC ATC CTG GGC ATG GGC TAC AAG GAG AAC GTG GAC	984					
Arg Ala Pro Glu Val Ile Leu Gly Met Gly Tyr Lys Glu Asn Val Asp						
195	200	205				
ATA TGG TCT GTG GGC TGC ATC ATG GGA GAA ATG GTT CGT CAC AAA ATC	1032					
Ile Trp Ser Val Gly Cys Ile Met Gly Glu Met Val Arg His Lys Ile						
210	215	220				

CTC TTT CCC GGA AGG GAC TAT ATT GAC CAG TGG AAC AAA GTC ATA GAG Leu Phe Pro Gly Arg Asp Tyr Ile Asp Gln Trp Asn Lys Val Ile Glu 225 230 235	1080
CAG CTA GGA ACT CCG TGT CCA GAA TTC ATG AAG AAA TTG CAG CCC ACC Gln Leu Gly Thr Pro Cys Pro Glu Phe Met Lys Lys Leu Gln Pro Thr 240 245 250 255	1128
GTC AGA AAC TAC GTG GAG AAC CGG CCC AAG TAT GCA GGC CTC ACC TTC Val Arg Asn Tyr Val Glu Asn Arg Pro Lys Tyr Ala Gly Leu Thr Phe 260 265 270	1176
CCC AAG CTC TTT CCA GAT TCC CTC TTC CCA GCG GAT TCC GAG CAC AAT Pro Lys Leu Phe Pro Asp Ser Leu Phe Pro Ala Asp Ser Glu His Asn 275 280 285	1224
AAA CTT AAA GCC AGC CAA GCC AGG GAC TTG TTG TCA AAG ATG TTA GTG Lys Leu Lys Ala Ser Gln Ala Arg Asp Leu Leu Ser Lys Met Leu Val 290 295 300	1272
ATT GAC CCA GCG AAG AGG ATA TCG GTG GAT GAC GCA TTG CAG CAT CCG Ile Asp Pro Ala Lys Arg Ile Ser Val Asp Asp Ala Leu Gln His Pro 305 310 315	1320
TAC ATC AAC GTT TGG TAC GAC CCT GCT GAA GTG GAG GCG CCT CCG CCT Tyr Ile Asn Val Trp Tyr Asp Pro Ala Glu Val Glu Ala Pro Pro 320 325 330 335	1368
CAG ATA TAT GAC AAG CAA TTG GAT GAA AGG GAG CAC ACC ATC GAA GAA Gln Ile Tyr Asp Lys Gln Leu Asp Glu Arg Glu His Thr Ile Glu Glu 340 345 350	1416
TGG AAA GAA CTC ATC TAC AAG GAA GTA ATG AAC TCA GAA GAG AAG ACT Trp Lys Glu Leu Ile Tyr Lys Glu Val Met Asn Ser Glu Glu Lys Thr 355 360 365	1464
AAG AAC GGC GTA GTC AAA GGC CAG CCC TCA CCT TCA GGT GCA GCA GTG Lys Asn Gly Val Val Lys Gly Gln Pro Ser Pro Ser Gly Ala Ala Val 370 375 380	1512
AAC AGC AGT GAG AGT CTC CCT CCA TCC TCA TCT GTC AAC GAC ATC TCC Asn Ser Ser Glu Ser Leu Pro Pro Ser Ser Ser Val Asn Asp Ile Ser 385 390 395	1560
TCC ATG TCC ACC GAC CAG ACC CTC GCA TCC GAC ACT GAC AGC AGC CTG Ser Met Ser Thr Asp Gln Thr Leu Ala Ser Asp Thr Asp Ser Ser Leu 400 405 410 415	1608
GAA GCC TCG GCG GGA CCG CTG GGT TGT TGC AGG TGACTAGCCG CCTGCCTGCG Glu Ala Ser Ala Gly Pro Leu Gly Cys Cys Arg 420 425	1661
AAACCCAGCG TTCTTCAGGA GATGACGCCA TGATAGAACA CAGCGCACAT GCACACACAC AGAGCTTGTA CACACACACA CACACACACA CACACGCACG CACGCACGCA CGCAAGCACG CACGCACGCA CAAATGCACT CACGCAATGT CAAGAAAAAA AAAAGTAGCG AGAGAGAGCG AGAGAGCCAA CGTAAACTA AGTTAAATCT TTCTGCGTGC TTCTCCAGAG TTCTGTATCG CAGCTGAGCT GAAATGTATA CTTAACTTCT AGTCGCGCTC GCTCGACTTT GGTCTCCCTC CGGCAGTGCT TACT	1721 1781 1841 1901 1961 1975



(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 426 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Ser	Lys	Ser	Lys	Val	Asp	Asn	Gln	Phe	Tyr	Ser	Val	Glu	Val	Gly	
1				5					10					15		
Asp	Ser	Thr	Phe	Thr	Val	Leu	Lys	Arg	Tyr	Gln	Asn	Leu	Lys	Pro	Ile	
			20					25					30			
Gly	Ser	Gly	Ala	Gln	Gly	Ile	Val	Cys	Ala	Ala	Tyr	Asp	Ala	Val	Leu	
		35					40					45				
Asp	Arg	Asn	Val	Ala	Ile	Lys	Lys	Leu	Ser	Arg	Pro	Phe	Gln	Asn	Gln	
	50					55					60					
Thr	His	Ala	Lys	Arg	Ala	Tyr	Arg	Glu	Leu	Val	Leu	Met	Lys	Cys	Val	
	65				70					75					80	
Asn	His	Lys	Asn	Ile	Ile	Ser	Leu	Leu	Asn	Val	Phe	Thr	Pro	Gln	Lys	
				85					90					95		
Thr	Leu	Glu	Glu	Phe	Gln	Asp	Val	Tyr	Leu	Val	Met	Glu	Leu	Met	Asp	
		100						105					110			
Ala	Asn	Leu	Cys	Gln	Val	Ile	Gln	Met	Glu	Leu	Asp	His	Glu	Arg	Met	
		115					120					125				
Ser	Tyr	Leu	Leu	Tyr	Gln	Met	Leu	Ser	Ala	Ile	Lys	His	Leu	His	Ser	
	130					135					140					
Ala	Gly	Ile	Ile	His	Arg	Asp	Leu	Lys	Pro	Ser	Asn	Ile	Val	Val	Lys	
	145				150					155					160	
Ser	Asp	Cys	Thr	Leu	Lys	Ile	Leu	Asp	Phe	Gly	Leu	Ala	Arg	Thr	Ala	
			165						170					175		
Gly	Thr	Ser	Phe	Met	Met	Thr	Pro	Tyr	Val	Val	Thr	Arg	Tyr	Tyr	Arg	
			180					185					190			
Ala	Pro	Glu	Val	Ile	Leu	Gly	Met	Gly	Tyr	Lys	Glu	Asn	Val	Asp	Ile	
		195					200					205				
Trp	Ser	Val	Gly	Cys	Ile	Met	Gly	Glu	Met	Val	Arg	His	Lys	Ile	Leu	
	210				215						220					
Phe	Pro	Gly	Arg	Asp	Tyr	Ile	Asp	Gln	Trp	Asn	Lys	Val	Ile	Glu	Gln	
	225				230					235					240	
Leu	Gly	Thr	Pro	Cys	Pro	Glu	Phe	Met	Lys	Lys	Leu	Gln	Pro	Thr	Val	
			245						250					255		
Arg	Asn	Tyr	Val	Glu	Asn	Arg	Pro	Lys	Tyr	Ala	Gly	Leu	Thr	Phe	Pro	
			260					265					270			
Lys	Leu	Phe	Pro	Asp	Ser	Leu	Phe	Pro	Ala	Asp	Ser	Glu	His	Asn	Lys	
		275					280					285				
Leu	Lys	Ala	Ser	Gln	Ala	Arg	Asp	Leu	Leu	Ser	Lys	Met	Leu	Val	Ile	
	290					295					300					
Asp	Pro	Ala	Lys	Arg	Ile	Ser	Val	Asp	Asp	Ala	Leu	Gln	His	Pro	Tyr	
	305				310					315					320	
Ile	Asn	Val	Trp	Tyr	Asp	Pro	Ala	Glu	Val	Glu	Ala	Pro	Pro	Pro	Gln	
			325						330					335		
Ile	Tyr	Asp	Lys	Gln	Leu	Asp	Glu	Arg	Glu	His	Thr	Ile	Glu	Glu	Trp	
		340					345						350			
Lys	Glu	Leu	Ile	Tyr	Lys	Glu	Val	Met	Asn	Ser	Glu	Glu	Lys	Thr	Lys	
		355					360					365				
Asn	Gly	Val	Val	Lys	Gly	Gln	Pro	Ser	Pro	Ser	Gly	Ala	Ala	Val	Asn	
	370					375					380					

Ser Ser Glu Ser Leu Pro Pro Ser Ser Ser Val Asn Asp Ile Ser Ser  
 385 390 395 400  
 Met Ser Thr Asp Gln Thr Leu Ala Ser Asp Thr Asp Ser Ser Leu Glu  
 405 410 415  
 Ala Ser Ala Gly Pro Leu Gly Cys Cys Arg  
 420 425

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2522 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 310...1575

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGGGCTTGAG TGAGCTAAAG ATTGGGTCTT CTTGGAAATC ACCTGTCTGT TATTATTTTT	60
AAACAATCGC TACACCTCCA AAGACTCTGC TCCTTACTCC GGTTTGGAAT GTGGCTAATG	120
ACTACCCAGT AGGGAGGATT TCTGGGGCAA ACAGCCGGAC CAGGATCCTA GTTCTCAGGC	180
ACGGAATGGC TATTGTGAGA ACAGACCAG CAGGATCATC GCAGATCTTG GTTATGGCCA	240
CTCAGGCAAG ACGCTGTTGA GTTAAGACCC CTTTCCCATG GATGAGAAGC CACAGAAGCA	300
GTGGTATTT ATG AGC CTC CAT TTC TTA TAC TAC TGC AGT GAA CCA ACC TTG	351
Met Ser Leu His Phe Leu Tyr Tyr Cys Ser Glu Pro Thr Leu	
1 5 10	
GAT GTG AAA ATT GCC TTT TGT CAG GGA TTC GAT AAA CAC GTG GAT GTG	399
Asp Val Lys Ile Ala Phe Cys Gln Gly Phe Asp Lys His Val Asp Val	
15 20 25 30	
TCA TCT ATT GCC AAA CAT TAC AAC ATG AGC AAA AGC AAG GTG GAC AAC	447
Ser Ser Ile Ala Lys His Tyr Asn Met Ser Lys Ser Lys Val Asp Asn	
35 40 45	
CAG TTC TAC AGT GTG GAA GTG GGG GAC TCA ACC TTC ACC GTT CTT AAG	495
Gln Phe Tyr Ser Val Glu Val Gly Asp Ser Thr Phe Thr Val Leu Lys	
50 55 60	
CGC TAC CAG AAC CTG AAG CCA ATT GGC TCT GGG GCT CAG GGA ATA GTC	543
Arg Tyr Gln Asn Leu Lys Pro Ile Gly Ser Gly Ala Gln Gly Ile Val	
65 70 75	
TGT GCT GCG TAC GAC GCT GTC CTT GAC AGA AAT GTG GCC ATT AAG AAG	591
Cys Ala Ala Tyr Asp Ala Val Leu Asp Arg Asn Val Ala Ile Lys Lys	
80 85 90	
CTC AGC AGA CCC TTC CAG AAC CAA ACT CAC GCC AAG AGG GCT TAC CGG	639
Leu Ser Arg Pro Phe Gln Asn Gln Thr His Ala Lys Arg Ala Tyr Arg	
95 100 105 110	
GAG CTG GTG CTC ATG AAG TGT GTG AAC CAT AAA AAC ATT ATT AGC TTA	687
Glu Leu Val Leu Met Lys Cys Val Asn His Lys Asn Ile Ile Ser Leu	
115 120 125	

TTA AAT GTT TTT ACA CCC CAG AAA ACG CTG GAG GAG TTC CAA GAT GTC Leu Asn Val Phe Thr Pro Gln Lys Thr Leu Glu Glu Phe Gln Asp Val 130 135 140	735
TAC TTA GTG ATG GAA CTG ATG GAC GCC AAC CTG TGT CAG GTG ATT CAG Tyr Leu Val Met Glu Leu Met Asp Ala Asn Leu Cys Gln Val Ile Gln 145 150 155	783
ATG GAG CTG GAC CAC GAG CGG ATG TCT TAC TTG CTG TAC CAG ATG CTG Met Glu Leu Asp His Glu Arg Met Ser Tyr Leu Leu Tyr Gln Met Leu 160 165 170	831
TGT GGC ATC AAG CAC CTC CAC TCC GCT GGG ATC ATC CAC AGG GAC TTA Cys Gly Ile Lys His Leu His Ser Ala Gly Ile Ile His Arg Asp Leu 175 180 185 190	879
AAA CCC AGT AAC ATT GTA GTC AAG TCT GAT TGC ACA CTG AAA ATC CTC Lys Pro Ser Asn Ile Val Val Lys Ser Asp Cys Thr Leu Lys Ile Leu 195 200 205	927
GAC TTC GGA CTG GCC AGG ACA GCG GGT ACA AGC TTC ATG ATG ACT CCG Asp Phe Gly Leu Ala Arg Thr Ala Gly Thr Ser Phe Met Met Thr Pro 210 215 220	975
TAT GTG GTG ACG CGA TAT TAC AGA GCC CCT GAG GTC ATC CTG GGC ATG Tyr Val Val Thr Arg Tyr Tyr Arg Ala Pro Glu Val Ile Leu Gly Met 225 230 235	1023
GGC TAC AAG GAG AAC GTG GAC ATA TGG TCT GTG GGA TGC ATC ATG GGA Gly Tyr Lys Glu Asn Val Asp Ile Trp Ser Val Gly Cys Ile Met Gly 240 245 250	1071
GAA ATG GTT CGC CAC AAA ATC CTC TTT CCC GGA AGG AGC TAT ATT GAC Glu Met Val Arg His Lys Ile Leu Phe Pro Gly Arg Ser Tyr Ile Asp 255 260 265 270	1119
CAG TGG AAC AAA GTC ATC GAG CAG CTA GGA ACT CCG TGT CCA GAG TTC Gln Trp Asn Lys Val Ile Glu Gln Leu Gly Thr Pro Cys Pro Glu Phe 275 280 285	1167
ATG AAG AAA TTG CAG CCC ACA GTC AGA AAC TAC GTG GAG AAT CGG CCC Met Lys Lys Leu Gln Pro Thr Val Arg Asn Tyr Val Glu Asn Arg Pro 290 295 300	1215
AAG TAC GCA GGA CTC ACC TTC CCC AAG CTC TTT CCA GAT TCC CTC TTC Lys Tyr Ala Gly Leu Thr Phe Pro Lys Leu Phe Pro Asp Ser Leu Phe 305 310 315	1263
CCA GCG GAT TCT GAG CAC AAT AAA CTT AAA GCC AGC CAA GCC AGG GAT Pro Ala Asp Ser Glu His Asn Lys Leu Lys Ala Ser Gln Ala Arg Asp 320 325 330	1311
TTG TTG TCT AAG ATG TTA GTG ATT GAC CCA GTG AAG AGG ATA TCG GTG Leu Leu Ser Lys Met Leu Val Ile Asp Pro Val Lys Arg Ile Ser Val 335 340 345 350	1359
GAC GAC GCA CTG CAG CAT CCG TAC ATC AAC GTT TGG TAC GAC CCG GCT Asp Asp Ala Leu Gln His Pro Tyr Ile Asn Val Trp Tyr Asp Pro Ala 355 360 365	1407

GAA GTG GAG GCG CCT CCG CCT CAG ATA TAT GAT AAG CAG CTG GAT GAA	1455
Glu Val Glu Ala Pro Pro Pro Gln Ile Tyr Asp Lys Gln Leu Asp Glu	
370 375 380	
AGG GAG CAC ACC ATC GAA GAA TGG AAA GAA CTT ATC TAC AAG GAG GTA	1503
Arg Glu His Thr Ile Glu Glu Trp Lys Glu Leu Ile Tyr Lys Glu Val	
385 390 395	
ATG AAC TCA GAA GAG AAG ACT AAG AAT GGC GTA GTC AAA AGC CAG CCC	1551
Met Asn Ser Glu Glu Lys Thr Lys Asn Gly Val Val Lys Ser Gln Pro	
400 405 410	
TCG CCT TCA GCA CAG GTG CAG CAG TGAACAGCAG TGAGAGTCTC CCTCCATCCT	1605
Ser Pro Ser Ala Gln Val Gln Gln	
415 420	
CGGCTGTCAA CGACATCTCC TCCATGTCCA CCGACCAGAC CCTCGCATCT GACACTGACA	1665
GCAGCCTGGA GGCCTCGGCG GGACCGTTGG GTTGTTCAG GTGACTAGCC GCCTGCCTGC	1725
GAAACCCAGC GTTCTTCAGG AGATGACGCG ATAGAACACA GCACACATGC ACACACACAG	1785
CTTGCTCTCA CACACACTCA GCTTGCTCAC ACACACACAC ACACATACAC ACAACACAC	1845
ACTGTCTCTC TCTCACACAC ACACACTGTC ACAACGCACT CACGAAAGGT CAAGAAAAA	1905
ATAACAATAG AGAGATCCAA CATAAAATTA AGTTAAATTT TTCTGCGTGC TTCTCCAAAG	1965
TTCTGTATCA CAGCTGAGCT GAAATGTATA CTTAACCTCT AGTTGCGCTC GCTTTGGTTT	2025
CCCTCCAGCA GTGCTTACTA CACAAGACAA ATCAGACACA ATTAGAGAAA CCTTCCCTA	2085
AAGTGTAAC TAAGTGGCTG CAGAACCAGC AACCTGTAAC TGCCCTTCAA ATGGCATGAG	2145
GAGGTGGGCA CGGGTCCGCG CCAGCATGTG TGTGTCTCTA TCTCGCGTCT ACCTGCTCTT	2205
CCGGCCTAGT CAGATGGATG TAGATACAGA TCCCGCATGT GTCTGTATT CAAACAGCACT	2265
TAGAGATGCT CCTGTCACTG TCCTCCAGGC TCCACCAAGA CACACACCGG GGTACCACAT	2325
GGTCCATTTT ATGTGATCTA TTAATCTGAC ATAAATCCAT CTGTAATATA TTGCCAGTAT	2385
ATAAGCTGTT TAGTTTGTTA ATTGCTTAAG CTGTATGTCT TATAAGAGAC TATGTAAAGG	2445
GGGAAATGG AGGCGTGAAC TCTCAGACCC TTGAAGATGT AGCTTCCGAA TTTGACCGTT	2505
AAATGGCACC GTATACC	2522

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 422 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met	Ser	Leu	His	Phe	Leu	Tyr	Tyr	Cys	Ser	Glu	Pro	Thr	Leu	Asp	Val
1				5					10					15	
Lys	Ile	Ala	Phe	Cys	Gln	Gly	Phe	Asp	Lys	His	Val	Asp	Val	Ser	Ser
		20						25					30		
Ile	Ala	Lys	His	Tyr	Asn	Met	Ser	Lys	Ser	Lys	Val	Asp	Asn	Gln	Phe
		35					40					45			
Tyr	Ser	Val	Glu	Val	Gly	Asp	Ser	Thr	Phe	Thr	Val	Leu	Lys	Arg	Tyr
	50					55					60				
Gln	Asn	Leu	Lys	Pro	Ile	Gly	Ser	Gly	Ala	Gln	Gly	Ile	Val	Cys	Ala
	65				70					75				80	
Ala	Tyr	Asp	Ala	Val	Leu	Asp	Arg	Asn	Val	Ala	Ile	Lys	Lys	Leu	Ser
			85						90				95		
Arg	Pro	Phe	Gln	Asn	Gln	Thr	His	Ala	Lys	Arg	Ala	Tyr	Arg	Glu	Leu
		100						105				110			
Val	Leu	Met	Lys	Cys	Val	Asn	His	Lys	Asn	Ile	Ile	Ser	Leu	Leu	Asn
		115					120					125			

Val Phe Thr Pro Gln Lys Thr Leu Glu Glu Phe Gln Asp Val Tyr Leu  
 130 135 140  
 Val Met Glu Leu Met Asp Ala Asn Leu Cys Gln Val Ile Gln Met Glu  
 145 150 155 160  
 Leu Asp His Glu Arg Met Ser Tyr Leu Leu Tyr Gln Met Leu Cys Gly  
 165 170 175  
 Ile Lys His Leu His Ser Ala Gly Ile Ile His Arg Asp Leu Lys Pro  
 180 185 190  
 Ser Asn Ile Val Val Lys Ser Asp Cys Thr Leu Lys Ile Leu Asp Phe  
 195 200 205  
 Gly Leu Ala Arg Thr Ala Gly Thr Ser Phe Met Met Thr Pro Tyr Val  
 210 215 220  
 Val Thr Arg Tyr Tyr Arg Ala Pro Glu Val Ile Leu Gly Met Gly Tyr  
 225 230 235 240  
 Lys Glu Asn Val Asp Ile Trp Ser Val Gly Cys Ile Met Gly Glu Met  
 245 250 255  
 Val Arg His Lys Ile Leu Phe Pro Gly Arg Ser Tyr Ile Asp Gln Trp  
 260 265 270  
 Asn Lys Val Ile Glu Gln Leu Gly Thr Pro Cys Pro Glu Phe Met Lys  
 275 280 285  
 Lys Leu Gln Pro Thr Val Arg Asn Tyr Val Glu Asn Arg Pro Lys Tyr  
 290 295 300  
 Ala Gly Leu Thr Phe Pro Lys Leu Phe Pro Asp Ser Leu Phe Pro Ala  
 305 310 315 320  
 Asp Ser Glu His Asn Lys Leu Lys Ala Ser Gln Ala Arg Asp Leu Leu  
 325 330 335  
 Ser Lys Met Leu Val Ile Asp Pro Val Lys Arg Ile Ser Val Asp Asp  
 340 345 350  
 Ala Leu Gln His Pro Tyr Ile Asn Val Trp Tyr Asp Pro Ala Glu Val  
 355 360 365  
 Glu Ala Pro Pro Pro Gln Ile Tyr Asp Lys Gln Leu Asp Glu Arg Glu  
 370 375 380  
 His Thr Ile Glu Glu Trp Lys Glu Leu Ile Tyr Lys Glu Val Met Asn  
 385 390 395 400  
 Ser Glu Glu Lys Thr Lys Asn Gly Val Val Lys Ser Gln Pro Ser Pro  
 405 410 415  
 Ser Ala Gln Val Gln Gln  
 420

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2347 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 615...1616

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGCAACTTTC	CTGACCCAGA	GGACCGGTAA	CAAGTGGCCG	GGAGCAACTT	TTGCAAATCT	60
CTTCTGCGCC	TTAAGGCTGC	CACCGAGACT	GTAAAGAAAA	GGAGAAGAGG	AACCTATACT	120
CATACCAAGT	CGCACAGGCC	TAAGTTGGGC	GAGGCCTAGC	CGCGGCTGCC	TAGCGTCCCC	180
CCCCCCTCA	CAGCGGAGGA	GGGGACAGTT	GTTGGAGGCC	GGGCGGCAGA	CCCGATCGCG	240
GGCCTCCACC	GAGAATCCG	TGACGACTGG	TCAGCACCGC	CGGAGAGCCG	CTGTTGCTGG	300
GACTGGTCTG	CGGGCTCCAA	GGAACCGCTG	CTCCCCGAGA	GCGCTCCGTG	AGTGACCGCG	360

ACTTTTCAAA	GCTCGGCATC	GCGCGGAGTC	CTACCAACGT	GAGTGCTAGC	GGAGTCTTAA	420										
CCCTGCGCTC	CCTGGAGCGA	ACTGGGGAGG	AGGGCTCAGG	GGGAAGCACT	CCCGTCTGGA	480										
GCGCACGCTC	TAAACAAACT	TTGTTACAGA	AGCAGGGACG	CGCGGGTATC	CCCCCGCTTC	540										
CCGGCGCGCT	GTTGCGGCCC	CGAAACTTCT	GCGCACAGCC	CAGGCTAACC	CCGCGTGAAG	600										
TGACGGACCG	TTCT	ATG	ACT	GCA	AAG	ATG	GAA	ACG	ACC	TTC	TAC	GAC	GAT	650		
	Met	Thr	Ala	Lys	Met	Glu	Thr	Thr	Phe	Tyr	Asp	Asp				
	1					5					10					
GCC	CTC	AAC	GCC	TCG	TTC	CTC	CAG	TCC	GAG	AGC	GGT	GCC	TAC	GGC	TAC	698
Ala	Leu	Asn	Ala	Ser	Phe	Leu	Gln	Ser	Glu	Ser	Gly	Ala	Tyr	Gly	Tyr	
	15						20					25				
AGT	AAC	CCT	AAG	ATC	CTA	AAA	CAG	AGC	ATG	ACC	TTG	AAC	CTG	GCC	GAC	746
Ser	Asn	Pro	Lys	Ile	Leu	Lys	Gln	Ser	Met	Thr	Leu	Asn	Leu	Ala	Asp	
	30					35					40					
CCG	GTG	GGC	AGT	CTG	AAG	CCG	CAC	CTC	CGC	GCC	AAG	AAC	TCG	GAC	CTT	794
Pro	Val	Gly	Ser	Leu	Lys	Pro	His	Leu	Arg	Ala	Lys	Asn	Ser	Asp	Leu	
	45				50					55					60	
CTC	ACG	TCG	CCC	GAC	GTC	GGG	CTG	CTC	AAG	CTG	GCG	TCG	CCG	GAG	CTG	842
Leu	Thr	Ser	Pro	Asp	Val	Gly	Leu	Leu	Lys	Leu	Ala	Ser	Pro	Glu	Leu	
				65					70					75		
GAG	CGC	CTG	ATC	ATC	CAG	TCC	AGC	AAT	GGG	CAC	ATC	ACC	ACT	ACA	CCG	890
Glu	Arg	Leu	Ile	Ile	Gln	Ser	Ser	Asn	Gly	His	Ile	Thr	Thr	Thr	Pro	
			80					85					90			
ACC	CCC	ACC	CAG	TTC	TTG	TGC	CCC	AAG	AAC	GTG	ACC	GAC	GAG	CAG	GAG	938
Thr	Pro	Thr	Gln	Phe	Leu	Cys	Pro	Lys	Asn	Val	Thr	Asp	Glu	Gln	Glu	
		95					100					105				
GGC	TTC	GCC	GAG	GGC	TTC	GTG	CGC	GCC	CTG	GCT	GAA	CTG	CAT	AGC	CAG	986
Gly	Phe	Ala	Glu	Gly	Phe	Val	Arg	Ala	Leu	Ala	Glu	Leu	His	Ser	Gln	
	110					115					120					
AAC	ACG	CTT	CCC	AGT	GTC	ACC	TCC	GCG	GCA	CAG	CCG	GTC	AGC	GGG	GCG	1034
Asn	Thr	Leu	Pro	Ser	Val	Thr	Ser	Ala	Ala	Gln	Pro	Val	Ser	Gly	Ala	
	125				130					135					140	
GGC	ATG	GTG	GCT	CCC	GCG	GTG	GCC	TCA	GTA	GCA	GGC	GCT	GGC	GGC	GGT	1082
Gly	Met	Val	Ala	Pro	Ala	Val	Ala	Ser	Val	Ala	Gly	Ala	Gly	Gly	Gly	
				145					150					155		
GGT	GGC	TAC	AGC	GCC	AGC	CTG	CAC	AGT	GAG	CCT	CCG	GTC	TAC	GCC	AAC	1130
Gly	Gly	Tyr	Ser	Ala	Ser	Leu	His	Ser	Glu	Pro	Pro	Val	Tyr	Ala	Asn	
			160					165					170			
CTC	AGC	AAC	TTC	AAC	CCG	GGT	GCG	CTG	AGC	AGC	GGC	GGT	GGG	GCG	CCC	1178
Leu	Ser	Asn	Phe	Asn	Pro	Gly	Ala	Leu	Ser	Ser	Gly	Gly	Gly	Ala	Pro	
		175					180					185				
TCC	TAT	GGC	GCG	GCC	GGG	CTG	GCC	TTT	CCC	TCG	CAG	CCG	CAG	CAG	CAG	1226
Ser	Tyr	Gly	Ala	Ala	Gly	Leu	Ala	Phe	Pro	Ser	Gln	Pro	Gln	Gln	Gln	
	190					195					200					
CAG	CAG	CCG	CCT	CAG	CCG	CCG	CAC	CAC	TTG	CCC	CAA	CAG	ATC	CCG	GTG	1274
Gln	Gln	Pro	Pro	Gln	Pro	Pro	His	His	Leu	Pro	Gln	Gln	Ile	Pro	Val	
	205				210					215					220	

CAG CAC CCG CGG CTG CAA GCC CTG AAG GAA GAG CCG CAG ACC GTG CCG Gln His Pro Arg Leu Gln Ala Leu Lys Glu Glu Pro Gln Thr Val Pro 225 230 235	1322
GAG ATG CCG GGA GAG ACG CCG CCC CTG TCC CCT ATC GAC ATG GAG TCT Glu Met Pro Gly Glu Thr Pro Pro Leu Ser Pro Ile Asp Met Glu Ser 240 245 250	1370
CAG GAG CGG ATC AAG GCA GAG AGG AAG CGC ATG AGG AAC CGC ATT GCC Gln Glu Arg Ile Lys Ala Glu Arg Lys Arg Met Arg Asn Arg Ile Ala 255 260 265	1418
GCC TCC AAG TGC CGG AAA AGG AAG CTG GAG CGG ATC GCT CGG CTA GAG Ala Ser Lys Cys Arg Lys Arg Lys Leu Glu Arg Ile Ala Arg Leu Glu 270 275 280	1466
GAA AAA GTG AAA ACC TTG AAA GCG CAA AAC TCC GAG CTG GCA TCC ACG Glu Lys Val Lys Thr Leu Lys Ala Gln Asn Ser Glu Leu Ala Ser Thr 285 290 295 300	1514
GCC AAC ATG CTC AGG GAA CAG GTG GCA CAG CTT AAG CAG AAA GTC ATG Ala Asn Met Leu Arg Glu Gln Val Ala Gln Leu Lys Gln Lys Val Met 305 310 315	1562
AAC CAC GTT AAC AGT GGG TGC CAA CTC ATG CTA ACG CAG CAG TTG CAA Asn His Val Asn Ser Gly Cys Gln Leu Met Leu Thr Gln Gln Leu Gln 320 325 330	1610
ACG TTT TGAGAACAGA CTGTCAGGGC TGAGGGGCAA TGGAAGAAAA AAAATAACAG Thr Phe	1666
AGACAAACTT GAGAACTTGA CTGGTTGCGA CAGAGAAAAA AAAAGTGTCC GAGTACTGAA GCCAAGGGTA CACAAGATGG ACTGGGTTTCG GACTGACGGC GCGCCAGTG TGCTCTGGAG TGGGAAGGAC GTGGCGCGCC TGGCTTTGGC GTGGAGCCAG AGAGCAGGCC TATTGGCCGG CAGACTTTGC GGAGCGCTGT GCCGCGCGCG ACCAGAACGA TGGACTTTTC GTTAACATTG ACCAAGAACT GCATGGACCT AACATTTCGAT CTCATTTCAGT ATTAAAGGGG GGTGGGAGGG GTTACAAACT GCAATAGAGA CTGTAGATTG CTTCTGTAGT GCTCCTTAAC ACAAAGCAGG GAGGGCTGGG AAGGGGGGGA GGCTTGTAAG TGCCAGGCTA GACTGCAGAT GAACTCCCCT GGCCTGCCTC TCTCAACTGT GTATGTACAT ATATATTTTT TTTTAATTTG ATGAAAGCTG ATTACTGTCA ATAAACAGCT TCCTGCCTTT GTAAGTTATT CCATGTTTGT TTGTTTGGGT GTCCTGCCCA GTGTTTGTAATAAAGAGATT TGAAGCATTG TGAGTTTACC ATTTGTAATA AAGTATATAA TTTTTTTATG TTTTGTCTTCT GAAAATTTCC AGAAAGGATA TTTAAGAAAA TACAATAAAC TATTGAAAAG T	1726 1786 1846 1906 1966 2026 2086 2146 2206 2266 2326 2347

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 334 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Thr	Ala	Lys	Met	Glu	Thr	Thr	Phe	Tyr	Asp	Asp	Ala	Leu	Asn	Ala
1				5				10						15	
Ser	Phe	Leu	Gln	Ser	Glu	Ser	Gly	Ala	Tyr	Gly	Tyr	Ser	Asn	Pro	Lys
		20					25						30		

Ile Leu Lys Gln Ser Met Thr Leu Asn Leu Ala Asp Pro Val Gly Ser  
35 40 45  
Leu Lys Pro His Leu Arg Ala Lys Asn Ser Asp Leu Leu Thr Ser Pro  
50 55 60  
Asp Val Gly Leu Leu Lys Leu Ala Ser Pro Glu Leu Glu Arg Leu Ile  
65 70 75 80  
Ile Gln Ser Ser Asn Gly His Ile Thr Thr Thr Pro Thr Pro Thr Gln  
85 90 95  
Phe Leu Cys Pro Lys Asn Val Thr Asp Glu Gln Glu Gly Phe Ala Glu  
100 105 110  
Gly Phe Val Arg Ala Leu Ala Glu Leu His Ser Gln Asn Thr Leu Pro  
115 120 125  
Ser Val Thr Ser Ala Ala Gln Pro Val Ser Gly Ala Gly Met Val Ala  
130 135 140  
Pro Ala Val Ala Ser Val Ala Gly Ala Gly Gly Gly Gly Tyr Ser  
145 150 155 160  
Ala Ser Leu His Ser Glu Pro Pro Val Tyr Ala Asn Leu Ser Asn Phe  
165 170 175  
Asn Pro Gly Ala Leu Ser Ser Gly Gly Gly Ala Pro Ser Tyr Gly Ala  
180 185 190  
Ala Gly Leu Ala Phe Pro Ser Gln Pro Gln Gln Gln Gln Gln Pro Pro  
195 200 205  
Gln Pro Pro His His Leu Pro Gln Gln Ile Pro Val Gln His Pro Arg  
210 215 220  
Leu Gln Ala Leu Lys Glu Glu Pro Gln Thr Val Pro Glu Met Pro Gly  
225 230 235 240  
Glu Thr Pro Pro Leu Ser Pro Ile Asp Met Glu Ser Gln Glu Arg Ile  
245 250 255  
Lys Ala Glu Arg Lys Arg Met Arg Asn Arg Ile Ala Ala Ser Lys Cys  
260 265 270  
Arg Lys Arg Lys Leu Glu Arg Ile Ala Arg Leu Glu Glu Lys Val Lys  
275 280 285  
Thr Leu Lys Ala Gln Asn Ser Glu Leu Ala Ser Thr Ala Asn Met Leu  
290 295 300  
Arg Glu Gln Val Ala Gln Leu Lys Gln Lys Val Met Asn His Val Asn  
305 310 315 320  
Ser Gly Cys Gln Leu Met Leu Thr Gln Gln Leu Gln Thr Phe  
325 330

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3967 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATACCAGAGA	CTCAAAAAA	AAAAAAAAGT	TCCAGATTGC	TGGACAATGA	CCCGGGTCTC	60
ATCCCTTGAC	CCTGGGAACC	GGGTCCACAT	TGAATCAGGT	GCGAATGTTT	GCTCGCCTTC	120
TCTGCCTTTC	CCGCCTCCCC	TCCCCCGGCC	GCGGCCCCGG	TTCCCCCCTT	GCGCTGCACC	180
CTCAGAGTTG	GCTGCAGCCG	GCGAGCTGTT	CCCGTCAATC	CCTCCCTCCT	TTACACAGGA	240
TGTCCATATT	AGGACATCTG	CGTCAGCAGG	TTTCCACGGC	CGGTCCCTGT	TGTTCTGGGG	300
GGGGGACCAT	CTCCGAAATC	CTACACGCGG	AAGGTCTAGG	AGACCCCTTA	AGATCCCAAA	360
TGTGAACACT	CATAGGTGAA	AGATGTATGC	CAAGACGGGG	GTTGAAAGCC	TGGGGCGTAG	420
AGTTGACGAC	AGAGCGCCCG	CAGAGGGCCT	TGGGGCGCGC	TTCCCCCCCC	TTCCAGTTCC	480
GCCCAGTGAC	GTAGGAAGTC	CATCCATTCA	CAGCGCTTCT	ATAAAGGCGC	CAGCTGAGGC	540
GCCTACTACT	CCAACCGCGA	CTGCAGCGAG	CAACTGAGAA	GACTGGATAG	AGCCGGCGGT	600



TCCGCGAACG	AGCAGTGACC	GCGCTCCAC	CCAGCTCTGC	TCTGCAGCTC	CCACCAGTGT	660
CTACCCCTGG	ACCCCTTGCC	GGGCTTTCCC	CAAACTTCGA	CCATGATGTT	CTCGGGTTTC	720
AACGCCGACT	ACGAGGCGTC	ATCCTCCCGC	TGCAGTAGCG	CCTCCCGGCG	CGGGGACAGC	780
CTTTCTTACT	ACCATTTCCC	AGCCGACTCC	TTCTCCAGCA	TGGGCTCTCC	TGTCAACACA	840
CAGGTGAGTT	TGGCTTTGTG	TAGCCGCCAG	GTCCGCGCTG	AGGGTCGCGG	TGGAGGAGAC	900
ACTGGGGTGT	GACTCGCAGG	GGCGGGGGGG	TCTTCTTTTT	TCGCTCTGGA	GGGAGACTGG	960
CGCGGTGAGA	GCAGCCTTAG	CCTGGGAACC	CAGGACTTGT	CTGAGCGCGT	GCACACTTGT	1020
CATAGTAAGA	CTTAGTGACC	CCTTCCCGCG	CGGCAGGTTT	ATTCTGAGTG	GCCTGCCTGC	1080
ATTCTTCTCT	CGGCCGACTT	GTTTCTGAGA	TCAGCCGGGG	CCAACAAGTC	TCGAGCAAAG	1140
AGTCGCTAAC	TAGAGTTTGG	GAGGCGGCAA	ACCGCGGCAA	TCCCCCTCC	CGGGGCGAGC	1200
TGGAGCAGGG	AGGAGGGAGG	AGGGAGGAGG	GTGCTGCGGG	CGGGTGTGTA	AGGCAGTTTC	1260
ATTGATAAAA	AGCGAGTTCA	TTCTGGAGAC	TCCGGAGCAG	CGCCTGCGTC	AGCGCAGACG	1320
TCAGGGATAT	TTATAACAAA	CCCCCTTTCC	AGCGAGTGAT	GCCGAAGGGA	TAAACGGGAA	1380
GCAGCAGTAG	GATGGAGGAG	AAAGGCTGCG	CTGCGGAATT	CAAGGGAGGA	TATTGGGAGA	1440
GCTTTTATCT	COGATGAGGT	GCATACAGGA	AGACATAAGC	AGTCTCTGAC	CGGAATGCTT	1500
CTCTCTCCCT	GCTTCATGCG	ACACTAGGGC	CACTTGCTCC	ACCTGTGTCT	GGAACCTCCT	1560
CGCTCACCTC	CGCTTTCTCT	TTTTTGTTTT	GTTTCAGGAC	TTTTGCGCAG	ATCTGTCCGT	1620
CTCTAGTGCC	AACCTTTATCC	CCACGGTGAC	AGCCATCTCC	ACCAGCCCAG	ACCTGCAGTG	1680
GCTGGTGCAG	CCCACCTCTG	TCTCCTCCGT	GGCCCCATCG	CAGACCAGAG	CGCCCCATCC	1740
TTACGGACTC	CCCACCCAGT	CTGCTGGGGC	TTACGCCAGA	GCGGGAATGG	TGAAGACCGT	1800
GTCAGGAGGC	AGAGCGCAGA	GCATCGGCAG	AAGGGGCAAA	GTAGAGCAGG	TGAGCAGCGA	1860
TTCTGGACCT	TTGTGGGCTG	GGGGGGGGGG	GGGGGGCGGA	GACTGACGCA	GAGACCACAC	1920
AACAGAGAAG	GGACGCTACT	GACTGCACTT	CCTGACCAGG	AGCTGTGGCT	GCTAGCCCTT	1980
TCCCTCCCTT	GTCAGATTTT	GACAGTTGGA	CCCAAGACAA	ACTCTAGACA	GTTTCCCTGA	2040
CAGCTTCCTA	CTTCATTCTC	TAGCCGGGGA	GCTTCTTTGT	TCCCCTGCTA	AAGATCTCAC	2100
TTTAAATGCA	AATCACACTC	TGCCTGCCAA	CTGCAGGTTA	GAAAAACTGC	TTACCCGAGA	2160
GGTGCGGGTG	CTGTAGGAGC	CAGTTTCACT	GGGGTGACTG	AATGGAGGTG	ACACTAGACA	2220
ACCTTAATCG	AATGTTGGTC	CTTTTCTTCT	ATAGCTATGT	CCTGAAGAGG	AAGAGAAACG	2280
GAGAAATCCG	AGGGAACGGA	ATAAGATGGC	TGCAGCCAA	TGCCGGAATC	GGAGGAGGGA	2340
GCTGACAGAT	ACACTCCAAG	CGGTAGGTTG	AACCAGCTGC	TGCTCCTGAA	ACTTTATTAA	2400
AGTTGGAGCT	TGGGACTATG	GGCGCAGGGT	CCTTGAGCAT	GCCCGTGTCT	TATGCTTTCT	2460
TATATCTCTC	CCTATGCAGG	AGACAGATCA	ACTTGAAGAT	GAGAAGTCTG	CGTTGCAGAC	2520
TGAGATTGCC	AATCTGCTGA	AAGAGAAGGA	AAAACCTGGAG	TTTATTTTGG	CAGCCCACCG	2580
ACCTGCCTGC	AAGATCCCCG	ATGACCTTGG	CTTCCCAGAG	GAGATGTCTG	TGGCCTCCCT	2640
GGATTGACT	GGAGGTCTGC	CTGAGGCTTC	CACCCAGAG	TCTGAGGAGG	CCTTCAACCT	2700
GCCCCTTCTC	AACGACCCTG	AGCCCCAAGCC	ATCCTTGGAG	CCAGTCAAGA	GCATCAGCAA	2760
CGTGGAGCTG	AAGGCAGAAC	CCTTTGATGA	CTTCTTGTTC	CCGGCATCAT	CTAGGCCAG	2820
TGGCTCAGAG	ACCTCCCGCT	CTGTGCCAGA	TGTGGACCTG	TCCGGTTCCT	TCTATGCAGC	2880
AGACTGGGAG	CCTCTGCACA	GCAATTCCTT	GGGGATGGGG	CCCATGGTCA	CAGAGCTGGA	2940
GCCCCGTGTG	ACTCCCGTGG	TCACCTGTAC	TCCGGGCTGC	ACTACTTACA	CGTCTTCCCT	3000
TGTCTTCACC	TACCCCTGAAG	CTGACTCCTT	CCCAAGCTGT	GCCGCTGCCC	ACCGAAAGGG	3060
CAGCAGCAGC	AACGAGCCCT	CCTCCGACTC	CCTGAGCTCA	CCCACGCTGC	TGGCCCTGTG	3120
AGCAGTCAGA	GAAGGCAAGG	CAGCCGGCAT	CCAGACGTGC	CACTGCCCGA	GCTGGTGCAT	3180
TACAGAGAGG	AGAAACACGT	CTTCCCTCGA	AGGTTCCTCG	CGACCTAGGG	AGGACCTTAC	3240
CTGTTCTGTA	AACACACCAG	GCTGTGGGCC	TCAAGGACTT	GCAAGCATCC	ACATCTGGCC	3300
TCCAGTCCTC	ACCTCTTCCA	GAGATGTAGC	AAAAACAAAA	CAAAACAAAA	CAAAAAACCG	3360
CATGGAGTGT	GTTGTTCTTA	GTGACACCTG	AGAGCTGGTA	GTTAGTAGAG	CATGTGAGTC	3420
AAGGCCTGGT	CTGTGTCTCT	TTTCTCTTTC	TCCTTAGTTT	TCTCATAGCA	CTAACTAATC	3480
TGTTGGGTTT	ATTATTGGAA	TTAACCTGGT	GCTGGATTGT	ATCTAGTGCA	GCTGATTTTA	3540
ACAATACCTA	CTGTGTTTCT	GGCAATAGCG	TGTTCCAAAT	AGAAACGACC	AATATTAAAC	3600
TAAGAAAAGA	TAGGACTTTA	TTTTCCAGTA	GATAGAAATC	AATAGCTATA	TCCATGTACT	3660
GTAGTCCTTC	AGCGTCAATG	TTCAATTGTC	TGTTACTGAT	CATGCATTGT	CGAGGTGGTC	3720
TGAATGTTCT	GACATTAACA	GTTTTCCATG	AAAACGTTTT	TATTGTGTTT	TCAATTTATT	3780
TATTAAGATG	GATTCTCAGA	TATTTATATT	TTTTTTTCTT	TTTTTTCTAC	CCTGAGGTCT	3840
TTGCACATGT	GGAAAGTGAA	TTTGAATGAA	AAATTTTAAG	CATTGTTTGC	TTATTGTTCC	3900
AGGACATTGT	CAATAAAAGC	ATTTAAGTTG	AATGCGACCA	CCTTCTTGCT	CTCTTTATTC	3960
TCAGTTT						3967

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 380 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met	Met	Phe	Ser	Gly	Phe	Asn	Ala	Asp	Tyr	Glu	Ala	Ser	Ser	Ser	Arg
1				5					10					15	
Cys	Ser	Ser	Ala	Ser	Pro	Ala	Gly	Asp	Ser	Leu	Ser	Tyr	Tyr	His	Ser
			20					25					30		
Pro	Ala	Asp	Ser	Phe	Ser	Ser	Met	Gly	Ser	Pro	Val	Asn	Thr	Gln	Asp
		35					40					45			
Phe	Cys	Ala	Asp	Leu	Ser	Val	Ser	Ser	Ala	Asn	Phe	Ile	Pro	Thr	Val
	50					55					60				
Thr	Ala	Ile	Ser	Thr	Ser	Pro	Asp	Leu	Gln	Trp	Leu	Val	Gln	Pro	Thr
	65				70				75					80	
Leu	Val	Ser	Ser	Val	Ala	Pro	Ser	Gln	Thr	Arg	Ala	Pro	His	Pro	Tyr
			85					90					95		
Gly	Leu	Pro	Thr	Gln	Ser	Ala	Gly	Ala	Tyr	Ala	Arg	Ala	Gly	Met	Val
			100					105					110		
Lys	Thr	Val	Ser	Gly	Gly	Arg	Ala	Gln	Ser	Ile	Gly	Arg	Arg	Gly	Lys
		115					120					125			
Val	Glu	Gln	Leu	Ser	Pro	Glu	Glu	Glu	Lys	Arg	Arg	Ile	Arg	Arg	
	130					135				140					
Glu	Arg	Asn	Lys	Met	Ala	Ala	Ala	Lys	Cys	Arg	Asn	Arg	Arg	Arg	Glu
	145				150					155					160
Leu	Thr	Asp	Thr	Leu	Gln	Ala	Glu	Thr	Asp	Gln	Leu	Glu	Asp	Glu	Lys
			165						170					175	
Ser	Ala	Leu	Gln	Thr	Glu	Ile	Ala	Asn	Leu	Leu	Lys	Glu	Lys	Glu	Lys
			180					185					190		
Leu	Glu	Phe	Ile	Leu	Ala	Ala	His	Arg	Pro	Ala	Cys	Lys	Ile	Pro	Asp
		195					200					205			
Asp	Leu	Gly	Phe	Pro	Glu	Glu	Met	Ser	Val	Ala	Ser	Leu	Asp	Leu	Thr
	210					215					220				
Gly	Gly	Leu	Pro	Glu	Ala	Ser	Thr	Pro	Glu	Ser	Glu	Glu	Ala	Phe	Thr
	225				230					235				240	
Leu	Pro	Leu	Leu	Asn	Asp	Pro	Glu	Pro	Lys	Pro	Ser	Leu	Glu	Pro	Val
			245						250					255	
Lys	Ser	Ile	Ser	Asn	Val	Glu	Leu	Lys	Ala	Glu	Pro	Phe	Asp	Asp	Phe
		260						265					270		
Leu	Phe	Pro	Ala	Ser	Ser	Arg	Pro	Ser	Gly	Ser	Glu	Thr	Ser	Arg	Ser
		275					280					285			
Val	Pro	Asp	Val	Asp	Leu	Ser	Gly	Ser	Phe	Tyr	Ala	Ala	Asp	Trp	Glu
	290					295					300				
Pro	Leu	His	Ser	Asn	Ser	Leu	Gly	Met	Gly	Pro	Met	Val	Thr	Glu	Leu
	305				310					315				320	
Glu	Pro	Leu	Cys	Thr	Pro	Val	Val	Thr	Cys	Thr	Pro	Gly	Cys	Thr	Thr
			325						330					335	
Tyr	Thr	Ser	Ser	Phe	Val	Phe	Thr	Tyr	Pro	Glu	Ala	Asp	Ser	Phe	Pro
		340						345					350		
Ser	Cys	Ala	Ala	Ala	His	Arg	Lys	Gly	Ser	Ser	Ser	Asn	Glu	Pro	Ser
	355						360						365		
Ser	Asp	Ser	Leu	Ser	Ser	Pro	Thr	Leu	Leu	Ala	Leu				
	370					375					380				

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AAGAAATGGA GGCTCATAAA TACCACAGCT

30

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATTGGAAGAA GACCAAAGCA AGAGCAACTA

30

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TAAGTAAGTA GTGCTGTATG AATACAGACA

30

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TACTGGCAAT ATATTACAGA TGGGTTTATG

30

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GTGTGCAGCT TATGATGCTA TTCTTGAA

28

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGCGTCACCA CATACGGAGT CATC

24

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CTGGAGGAGT TCCAAGATGT CTACT

25

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TGGAAAGAGC TTGGGGAAGG TGAG

24